PORTFOLIO PERFORMANCE MEASUREMENT

BY

GILBERT W. BICKUM

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1987

TABLE OF CONTENTS

Page

ABSTRACT iv

CHAPTER

I INTRODUCTION 1

- Portfolio Performance Measurement . 1
- Scope of this Study 3
- Data 4
- Hypotheses 5

II REVIEW OF PRIOR RESEARCH 7

- Introduction 7
- Sharpe, Treynor and Jensen Models . 9
- Fama Dichotomization of Returns . 12
- Manager Comparisons 16
- Small Firm Effect 17
- Value Line Studies 18
- Contemporary Studies 26

III HYPOTHESES AND JUSTIFICATION 29

IV THE DATA 39

- Data Sources 39
- The Sample 41
- Issues Used in the Study 41

V METHODOLOGY 43

- The First Approach 44
- The Second Approach 49
- The Third Approach 58
- Other Tests 62
  - Sharpe, Treynor and Jensen Tests 63
  - Tests by Stated Fund Objective . 63
  - Test of Consistency 65

VI RESULTS 67

Introduction and Summary 67
  Results of Tests of Hypotheses . 67
  Discussion 69
Test Methodology One 70
  Results of Buy-Hold-Sell Tests . 70
  Discussion of Buy-Hold-Sell Test results 73
  Results of Individual Fund Tests 74
  Test for Differences in Risk . 76
Test Methodology Two 77
  Discussion 77
  Results of Test Methodology Two 78
  Discussion of Test Results . . 79
  Lagged Return Test 80
Test Methodology Three 81
  Discussion 81
  Test Three Results 82
Additional Tests 82
Summary of Tests 83
Test of Significance by Fund Objective 87
Tests of Consistency 87
Summation of Results 91
  Buy-Sell-Hold Tests 91
  Actual-Index Test 92
  Actual-Initial Holding Test . . 92
  Other Tests and Correlation . . 92

VII SUMMARY 94
  Final Summation 94
  Suggestions for Further Research . 94

APPENDIX STATISTICAL METHODOLOGY 98

BIBLIOGRAPHY 103

BIOGRAPHICAL SKETCH 113

Abstract of Dissertation Presented to the Graduate
School of the University of Florida in Partial Fulfillment
of the Requirements for the Degree of Doctor of Philosophy

PORTFOLIO PERFORMANCE MEASUREMENT

By

Gilbert W. Bickum

December 1987

Chairman: Dr. Robert C. Radcliffe
Major Department: Finance, Real Estate and Insurance

This study examines the "stock selection" abilities of 293 mutual funds during the 8 1/2 year period from 1 January, 1977, through 30 June, 1985. Three related tests are conducted. In the first test, returns on stocks held, purchased and sold by these funds during each of the 33 quarters are compared over the succeeding three, six and twelve months. In the second test, an index is created for each fund based on the fund's holdings over the test period and the returns on this index are compared with each fund's actual equity holding returns over the same test period. The third test examines returns on each fund's actual quarterly holdings with the returns which

would have been earned by the fund if it had simply continued to hold the initial equities observed at 3/31/77.

The results of these tests indicate that mutual fund managers as a whole seem to be able to select stocks that will perform better than the universe from which they select. The greater skill seems to be an ability to identify the stocks that will perform worse than this same universe.

CHAPTER I

INTRODUCTION

Portfolio Performance Measurement

The measurement of portfolio performance has been a major problem as long as there have been portfolios. Since Treynor, Sharpe and Jensen developed methods and models for portfolio evaluation, the interest in this topic has become more widespread and the models used in this measurement process have become more complex. This topic is important in a theoretical sense because it has obvious implications about market efficiency (Fama, 1970) and in a practical sense because investors clearly are interested in evaluating the performance of various managers.

A broad body of academic literature has sprung up, grown to maturity and spawned new generations of research concerning this controversial subject. Portfolio performance measurement is controversial because of the wide variance in research-backed opinion concerning the value of portfolio management. The consensus among theorists (primarily academic) appears to be that there is no discernable value to the portfolio management function, in that managed portfolios do not appear to outperform unmanaged portfolios to any statistically significant extent. On the other hand, the stated consensus among practicing portfolio managers is that active management does add value to portfolios of assets.

The bibliography lists many studies on each side of the issue and some that are in the middle. Not only is there disagreement concerning the results and general implications of the empirical studies listed, but in addition, many theorists take issue with the methodology.

At issue with current test methods are Levy (1972, 1981, 1982), Reinganum (1981) and Roll (1977, 1978). These researchers raise several objections to extant research, thereby helping to explain the diversity of results. Most of the objections stem from the use of the Capital Asset Pricing Model (CAPM) as a measurement tool. Research results are dependent on the choice of the benchmark portfolio that is used for comparison, on the choice of a proxy for the risk free rate and on the method used for beta estimation. The lack of stationarity and the large standard errors encountered when estimating market model betas add to the problem. In addition to CAPM problems, there can be differences in results due to the time period chosen and a lack of significance due to sample size limitations.

Scope of this Study

Recent literature in this area has suggested that performance measurement which is done by comparing individual portfolio results to a broadly based index might not be appropriate. Reinganum (1981), Rosenberg (1981), Ferguson (1980) and others have suggested that a manager's performance should be compared to that of a universe more nearly like the holdings of the manager being evaluated. The rationale for this argument is that it is inappropriate to include in a comparison securities which the manager does not intend to hold.[1]

In this study, we try to eliminate many of the theoretical and empirical difficulties which exist in current performance studies. We avoid the problems noted by Roll (1978) by eliminating the "market" proxy and the "risk free or zero beta" proxy. We also follow the

---

1. There are many reasons that managers exclude certain instruments from their available set. Some classes of portfolios have legal constraints; others are limited by firm policy. Most are limited simply by the managers' ability to assimilate information and a perception commonly expressed by managers that each manager has a particular "comparative advantage" in analyzing a particular type of security.

advice of the authors cited above by using each manager's relevant universe when making benchmark comparisons.

This study attempts to answer empirically the question of whether portfolio managers possess special skills or knowledge that allow them to choose stocks which produce superior returns. It is not our intention to identify the best or better portfolio managers, but to make definitive statements concerning the handling of a particular set of equities (that is, the stock purchase, sale and retention decisions of managers). This dissertation then, centers on the issues of market efficiency, information asymmetry and performance measurement.

Data

The vehicles chosen for this study are 293 mutual funds that have been in existence since 1 January, 1977. The decisions made by fund management are examined by calculating and comparing returns earned on fund equity holdings. Tests are conducted to compare the returns earned on stocks held, bought, and sold in a given time frame over various time periods. Tests are also done to compare the returns on stocks actually held to the returns earned on both average and initial portfolio holdings.

Hypotheses

$H0:_1$ There is no statistical difference between the returns earned on the stocks held by mutual fund(s) at the end of a given quarter and the stocks bought by mutual fund(s) during that same quarter over the next three, six and twelve months.

$H0:_2$ There is no statistical difference between the returns earned on the stocks held by mutual fund(s) at the end of a given quarter and the stocks sold by mutual fund(s) during that same quarter over the next three, six and twelve months.

$H0:_3$ There is no statistical difference between the returns earned on the stocks bought by mutual fund(s) during a given quarter and the stocks sold by mutual fund(s) during that same quarter over the next three, six and twelve months.

$H0:_4$ There is no statistical difference between the returns earned on the actual stock holdings of mutual funds over 33 quarters (1 January 1977 through 31 March 1985) and the returns over that same time period of a portfolio consisting of the dollar weighted average stock holdings of each manager.

$H0:_5$ There is no statistical difference between the returns earned on the actual stock holdings over 33 quarters and the returns that would have been earned on each fund's holdings as of the end of the first quarter.

This dissertation is organized into seven chapters. In Chapter I, the introduction, the author lays out the framework of the dissertation. In Chapter II, a survey of the literature bearing on this topic is presented. Chapter III contains the hypotheses. The justification of the hypotheses and their relationship to this study are also included in this chapter.

The data sources utilized in this study are listed in Chapter IV, and in Chapter V, the methodology used for the study is laid out. Chapter VI gives the results of the study. A summary, a conclusion and suggestions for further research are presented in Chapter VII.

CHAPTER II

REVIEW OF PRIOR RESEARCH

Introduction

As we proceed through this section, it will become apparent that theorists and practitioners are not in agreement with each other and that theorists do not agree among themselves. The closest we can come to a united front is among the practitioners who generally support the notion that portfolio managers outperform the market. This is not especially shocking, since this conclusion is consistent with the practitioners' children being fed on a regular basis.

Somewhat more surprising is the state of affairs among theoreticians and empirical researchers. Using various methods, this group has conducted literally hundreds of studies which conclude alternatively that portfolio managers do or do not earn excess returns.

Early papers in this field of research did not utilize risk measures and compared managers to each other or to an index. Several more recent articles also used return information only to test for consistency of performance across time (Brightman & Haslanger, 1980; Dunn & Theisen, 1983). These studies found no evidence of consistency or predictability and, in fact, found some negative correlation between performance in different time periods. (In the first study, the negative correlation was significant at the 5% level; in the second the negative correlation seemed to be dependent on the volatility of the market.) Other studies, using various methods which were not adjusted for risk, came to the opposite conclusion (Baesel, Shows & Thorpe, 1982; Lerner & Theerathorn, 1983).

The vast majority of research in this area has attempted to include an allowance for risk, the most common being some form of CAPM methodology. Most researchers used the CAPM in a strict sense, and many used a market model of the form

$$R_{it} = \alpha_i + \beta_i R_{mt} + \varepsilon_{it}$$

where R_{it} is the return on stock i in time t,

α_i is the regression intercept term,

β_i is the regression slope term,

R_{mt} is the market return in time t,

ε_{it} is a residual error term.

Sharpe, Treynor and Jensen Models

The methods and models developed by Sharpe, Treynor and Jensen have been used extensively and, thus, deserve more detailed attention at this point.

Jensen (1968) evaluated the performance of mutual funds from 1945 to 1964. In order to conduct this evaluation, he devised what has become known as the Jensen performance index (JPI) to indicate performance better (or worse) than would be expected for a given period. The JPI is alpha in the following time series regression.

[Eq 1] JPI: $R_{pt} - Rf_t = \alpha_p + \beta_{pk}(RI_t - Rf_t) + \mu_{pt}$

where

R_{pt} = the return on the portfolio of interest in time (t),

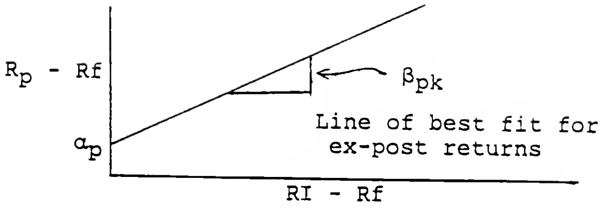
Rf_t = the risk free rate in time (t),

α_p = the regression intercept term,

β_{pk} = the portfolio Beta over the t(t=1->k) periods,

RI_t = the return on the comparison index in time t,

μ_{pt} = a residual term with mean zero and variance σ.



Jensen Measure

FIGURE ONE

A positive α_p indicates superior performance and a negative α_p indicates inferior performance by the portfolio in question in period t. A benefit of the JPI is that it will distinguish between superior, inferior, or index matching performance regardless of the risk (β) level. The JPI also allows the comparison of portfolios with different risk levels, but cannot distinguish between two index-superior or index-inferior portfolios with different risk levels.

Treynor (1965) suggested a measure of portfolio performance which is essentially the JPI normalized for systematic risk. The Treynor Performance Index (TPI) is given below:

[Eq 2] TPI: $(R_p - Rf)/\beta_{pk} = \alpha_p/\beta_{pk} + RI - Rf$

This index ignores the residual term and drops the t subscripts, but otherwise is identical to the JPI divided by β_{pk}. This index has the advantages of the JPI and also gives the excess (+,-) return per unit of systematic risk.

Sharpe (1966) devised the Sharpe Performance Index (SPI) which normalizes the JPI for total portfolio risk. (The Sharpe measure is appropriate only for aggregate portfolios and not for parts of the aggregate.) Sharpe divides the JPI by portfolio standard deviation to arrive at the model given below:

[Eq 3] SPI: $(R_p - Rf)/\sigma(R_p) = \alpha_p/\sigma(R_p) + (RI - Rf)/\sigma(RI)$

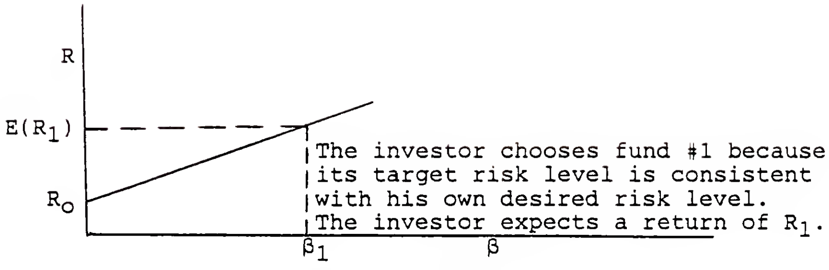
The SPI, therefore, gives excess (+,-) performance per unit of portfolio standard deviation.

The vast majority of risk adjusted studies evaluating or comparing portfolio management utilize one or more of these performance measures. Studies by Jensen (1969); Black, Jensen and Scholes (1972); and Friend and Blume (1970) used the Jensen measure to test for risk adjusted excess performance, finding that low beta funds had higher excess returns (a positive alpha) while higher beta funds generally underperformed the index (had a negative alpha). In 1974, McDonald, using the Sharpe, Treynor and Jensen measures came to the opposite conclusion. Shawky (1982) used the same measures and concluded that returns were generally commensurate with systematic risk. Other studies by Stock (1982) and Firth (1977) also found no excess returns (Firth actually found evidence of inferior performance). The common threads in the above cited works are that the methodology is similar (CAPM based) and that the results are similar in that no evidence of systematically superior performance was found. The differences in conclusions among these studies can easily be attributed to the choices of risk free assets, market proxies and time periods.

Fama Dichotomization of Returns

In 1972, Fama devised a method of breaking down the components of risk and return into several identifiable parts (investor's risk, manager's risk, selectivity, diversification). These components can be used to identify, illustrate and measure investment risk and manager/investor skills.

Portfolio risk (as proxied by portfolio beta) may be divided into the risk taken or desired by the investor and the risk taken by the portfolio manager in excess of or less than the investor's desired risk level. To illustrate this point, let us assume that an investor chooses a fund based upon a historical or expected fund risk level. This is what we call the fund's target risk level as indicated in the diagram below:

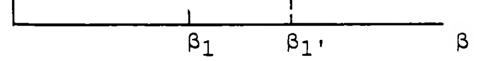


The investor chooses fund #1 because its target risk level is consistent with his own desired risk level. The investor expects a return of R_1.

Investors' Risk

FIGURE TWO

Suppose that in a given period the fund manager feels that the market will do better than normal (or the manager has special information). In this case, the fund manager might take on a higher than normal degree of risk as shown below:



The manager chooses a risk level of $\beta_{1'}$ which is higher than the investor's chosen risk level of β_1. The expected return is now $E(R_{1'})$ which is higher than $E(R_1)$.

Managers' Risk

FIGURE THREE

The total return may be divided into that expected due to the investor's chosen risk level (the fund target risk level) and the extra expected return due to the manager's taking on a temporarily higher risk level. This relationship is illustrated further in the model below.

[Eq 4] $E(R_1') = R_o + E(R_1 - R_o) + E(R_1' - R_1)$

where

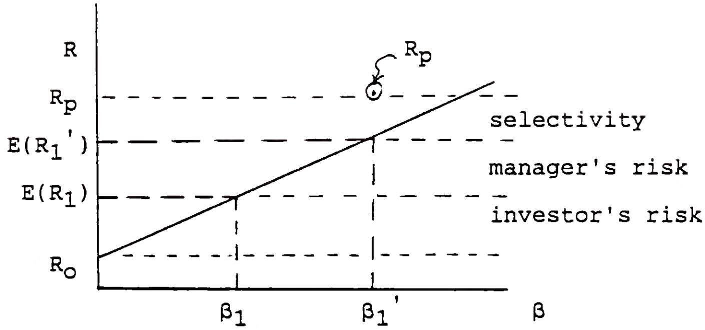
$E(R_1')$ is the total expected return,

R_o is the riskless rate of return,

$E(R_1 - R_o)$ is the return due to the investor's risk,

$E(R_1' - R_1)$ is the return due to the manager's risk.

Ex-post, the risk/return line might have any slope, and the actual portfolio return might be on, below, or above the ex-ante risk/return line. If the portfolio manager has made superior stock selection decisions prior to a given time period (or if he/she was lucky), the portfolio return point will fall above the risk/return line. If, on the other hand, the manager made poor stock selection decisions prior to a given time period (or if the manager was unlucky) the portfolio return point will fall below the risk/return line. Fama names this excess (+,-) return "selectivity." The case illustrated below adds positive selectivity to the previous diagram.



Selectivity Return

FIGURE FOUR

The portfolio return is shown as R_p in this period. This return may be divided into (1) the riskless return, (2) the return expected due to the risk level chosen by the investor (target risk level), (3) the return expected due to the added risk assumed by the portfolio manager (timing), and (4) the excess or unexpected return due to superior (or lucky) stock selection in this particular time period. These returns are illustrated algebraically in equation (5) below.

[Eq 5] $R_p = R_o + (R_1 - R_o) + (R_1' - R_1) + (R_p - R_1')$

where

R_p is the total return in this time period,

R_o is the riskless rate of return,

$(R_1 - R_o)$ is the return due to the investor's risk,

$(R_1' - R_1)$ is the return due to the manager's risk,

$(R_p - R_1')$ is the return due to stock selection.

By breaking down performance into components, Fama gave researchers a way of evaluating different apparent skills and attributing unusual performance to the presence (or absence) of these skills. The weakness in Fama's method is the use of CAPM based methodology, the same theoretical base that stands under the Sharpe, Treynor and Jensen measures.

Kon and Jen (1977) and Kon (1983) used the Fama methodology to examine mutual fund managements' selectivity and timing abilities. Both studies found evidence of positive selectivity, but, while the latter study found evidence of attempts at timing, there was no evidence of superior timing ability.

Manager Comparisons

Studies of portfolio managers and analysts were conducted by Bogle and Twardowski (1980) and by Groth, Lewellen, Schlarbaum and Lease (1978) showing that mutual funds and analysts tended to outperform other portfolio managers. The 1978 study tracked the performance of stocks after they had been recommended by analysts and the 1980 study compared different classes of portfolio managers (mutual funds, investment counselors, insurance companies, banks). The 1978 study compared the performance of analysts, finding that the more experienced analysts seemed to outperform less experienced ones. This study suffers from survivor bias and is weakened by measurements being taken in the month that buy-sell decisions were made. The 1980 study found that mutual fund managers outperformed other managers, with investment counselors second, insurance companies third and banks last. There was no adjustment for risk in this study.

Small Firm Effect

The next group of studies that are important antecedents to this one includes those falling under the broad heading of "the small firm effect." These studies actually include small firm, neglected firm, turn of the year and P/E ratio studies. These studies, conducted by Lustig and Leinbach (1983), Roll (1983a, 1983b) Edmister and Greene (1980), Goodman and Peavy (1983); Arbel, Carvel and Strebel (1983) and Arbel and Strebel (1983) all came to the same conclusion that, using current methodology, small firms (some would include neglected firms and firms with low P/E ratios in this catagory) significantly outperformed large firms on a risk adjusted basis. They found that the CAPM does not do a good job of describing the risk/return relationship for small firms and that total risk is more highly correlated with the returns on these firms than is systematic risk. Roll (1981) examined the methodology used in these studies for flaws and an explanation of the effect. He did find a few minor problems, but concluded that the excess returns remained even after correcting for the flaws. The P/E studies concluded that stocks with lower P/E ratios outperform those with higher P/E ratios, but P/Es have been shown to be highly correlated with size which is why they were included with this group. Neglected firms (defined as firms not followed and/or held by institutions) as well belong in this group for two reasons. First, size and "neglect" are highly correlated. Second, the same factors that may contribute to small firm premiums (lack of information, thin trading, not favored by institutions) most likely hold for neglected firms.

Value Line Studies

The next group of studies is especially relevant because of the close relationship to this one. The methodology is different, but the skills being examined are quite similar. As stated in the introduction, performance evaluation is a current and widely debated issue within the theory of finance. The vast majority of the studies in this area are based on methods devised by Jensen, Sharpe, Treynor and Fama as have been illustrated above in some detail. The theoretical basis for these methods was criticized by Roll (1977,1978) and supported by Mayers and Rice (1979). In addition, potential empirical problems were pointed out by Banz (1981), Litzenberger and Ramaswamy (1979), Reinganum (1981) and others. Solutions to these problems were suggested by Cornell (1979), Fama (1972), and Ferguson (1980).

Probably the most thoroughly investigated portfolio to the present time has been the rankings of the Value Line Investment Survey, the world's largest published advisory service.[1] Black (1971), using Jensen's CAPM based performance evaluation technique, found monotonically increasing returns on stocks ranked 5 to 1, in accordance with the Value Line rankings. In addition, he found statistically significant performance for portfolios 1, 2, 4 and 5, and economically significant performance for portfolios 1 and 5. These results are in conflict with at least the semi-strong tests of the Efficient Markets Hypothesis.

---

[1] Since April, 1965, Value Line has published weekly performance predictions, ranking approximately 1700 stocks. Stocks are ranked from 1 to 5 with 1 being most favorable and 5 being least favorable (3 is average or neutral). Approximately 53% of the stocks are rated as a 3, 18% are rated as a 2 or as a 4 and 6% are rated as a 1 or as a 5. Value Line relies heavily on past performance for its evaluations.

Because of the theoretical and empirical problems cited above, Copeland and Mayers (1982) tested the Value Line data over the same time period as Black using a variant of the methodology suggested by Cornell. Cornell (and others) had suggested that the CAPM/Benchmark problems could be largely eliminated by the following method.

(1) Using a sample prior to the test period, estimate the mean return on all individual securities that are held in the portfolio to be examined during the test period. In a competitive market the appropriate risk premium, which may not be measurable, will be impounded in the return.

(2) At the beginning of the test period, observe the weights for the portfolio of the individual under study and compute the expected return for the first interval using the returns computed in step one.

(3) Record the actual return for the first interval and compute the unexpected return.

(4) Proceed in this manner throughout the test period to compile a series of unexpected returns.

(5) Test to see if the mean of this series is significantly different from zero. (Note that as more observations on returns become available in moving through

the test period, the estimate of the expected return can be updated.)[2]

Copeland and Mayers did not follow Cornell's method exactly, but did agree that benchmark returns could be estimated best by examination of the data under study rather than from a disembodied market index and risk free rate. The most common way to estimate expected returns from the data would be (as Cornell suggested) the use of past return data. Copeland and Mayers observed that the use of a prior benchmark could well detract from the probability of finding statistical significance in the Value Line case, as Value Line uses historical price and return data to make its recommendations.[3] Therefore, if abnormally high returns in the estimation period were followed by abnormally high returns in the test period, no unusual performance would be noted. [Value Line (and others) purchase and recommend stocks after noting unusually high returns. Therefore, if recent historical returns are used as a benchmark, this benchmark would be biased upward. This tendency is also noted and tested in this study.]

---

[2] It should be noted that following this advice would weaken the results.

[3] In addition, past returns are susceptible to gaming. Test methods and benchmarks can be chosen that make test period results look good or bad at the option of the researcher.

While this is a problem in all event studies and studies which estimate return generating processes from a data set under study, Copeland and Mayers felt that the problem would be exaggerated by Value Line's use of return price/earnings and relative strength measures from periods just prior to the recommendations. For this reason, Copeland and Mayers skipped the 26 weeks following the announcements (the test period) and used the subsequent 26 or 52 weeks as benchmark periods.[4]

Hoping to increase the chances of finding significance, Copeland and Mayers employed this future benchmark technique for return comparison. They used mean return and market model $[R_{pt} = \alpha_p + \beta(Rm_t) + \varepsilon_{pt}]$ benchmarks. Copeland and Mayers obtained Value Line rankings for 24 periods of 26 weeks each from 26 November, 1965, through 3 February, 1978. They formed five portfolios to conform with the Value Line rankings in each period and calculated the weekly returns on the stocks held in each portfolio for the next 26 weeks using the University of Chicago's Center for Research in Securities Prices (CRSP) files. The portfolio returns

---

4. Obviously, there could be some contamination here since (1) Value Line's recommendations are ostensibly predictors of performance over the next year; (2) if Value Line has predicting ability that extends beyond 26 weeks, the results would be weakened; (3) a permanent change in the return generating process (an important occurrence) would not be detected.

were then calculated by taking an equally weighted return on stocks held in each portfolio [$R_{pt} = \ln(1 + \sum_{j=1}^{n} R_{jt}/N_t)$], where R_{jt} is the weekly rate of return for security j during week t. (Non-log returns were also calculated with similar results.)

Copeland and Mayers's results were similar to Black's, except that the absolute values of the differences were about half of those found by Black. They also found monotonically increasing returns from portfolios 5 to 1, but statistically significant performance was found only for portfolio 5.

In a separate test of the same data, Copeland and Mayers formed portfolios of stocks that were either upgraded or downgraded since the last recommendation.[5] They also tested the performance of these portfolios in the same manner as their other tests as well as the combined strategy of going long the "up" portfolio and short the "down" portfolio. The results of this set of tests were similar to those of the previous tests. The "up" portfolio significantly outperformed the "down" portfolio for the total test history (48 periods) and for three 16-week sub-periods (using the 52-week benchmark).[6]

---

[5] These portfolios would necessarily contain stocks currently rated 2, 3, 4 or 5 for the "down" portfolio and 1, 2, 3 or 4 for the "up" portfolio.

[6] For the 26-week benchmark, significance was limited to the whole history and the 1st 16-week sub-period.

What is remarkable about these results is that the significance was concentrated almost entirely in the "down" portfolio. In addition, while the "down" and the combined portfolios showed significance, the significance in the combined portfolio was attributed to the short or down side as well. These results are important in the light of the empirical findings of this dissertation. Copeland and Mayers (and Black before them) were testing the ability of portfolio managers to pick stocks from a given universe that would in some future period generate unusual returns. These unusual returns might be more or less than those of the universe as a whole. Managers might possess skills or have superior information that would allow them to ferret out those equities which will produce unusually high returns in the future (stocks that are "underpriced" or should be bought) or equities which will produce unusually low returns in the future (stocks which are "overpriced" or should be sold--perhaps short) or both.

Copeland and Mayers's results indicate that the Value Line analysts' "good" stocks can not be statistically differentiated from the probability distribution containing the relevant universe (all CRSP stocks), but that the analysts' "bad" stocks are drawn from a different distribution. More simply put, there may be indicators that presage corporate problems resulting in lowered earnings and price levels which are not picked up by the market in general, but are picked up by some analysts. At the same time, these analysts are not able to discern signals of firms with unusually good prospects.[7] For individuals possessing large diversified portfolios and/or for individuals who plan to transact for liquidity purposes anyway, this information might be of significant value. Certainly, if the managers of mutual funds (which are large diversified portfolios) could obtain information of this kind, their funds would benefit significantly. Not only do mutual funds have a large equity base to work from, but mutual funds (as well as other large institutional investors) have the added advantage of incurring lower transactions costs. The nature of mutual fund operations increases the value of this kind of forecasting as well. Mutual fund shares are continually being purchased and sold, necessitating frequent trading. If managers can pick stocks with average prospects and disinvest using stocks with below average prospects, they will "beat the market" simply by truncating the lower part of the return distribution.

---

[7] There is another force here which might be considered of prime importance to some investors. Copeland and Mayers assume 1% transactions costs and a .62% spread. It is well known that shorting transactions costs are far greater than transactions costs for long transactions. Thus, a strategy of shorting portfolio 5 or "down" portfolio stocks might not be profitable due to these higher costs despite the indicated significance.

Contemporary Studies

Most recently, Grinblatt (1986) and Grinblatt and Titman (1986) have developed a period weighting model to identify managers with superior information. This measure is, on average, zero for uninformed managers and positive for informed managers. This measure finds a time series of weights that sum to one and have the following properties.

(i) They are non-negative.

(ii) They make the weighted average of the time series of excess returns on a mean-variance efficient portfolio zero.

This method includes only those assets that can be held by the portfolio manager in the proxy portfolio. The returns on the stocks (assets) held by the manager under study during the test period are compared with the returns on these same stocks when not held (typically at some time prior to the test period). The difference in the weightings the manager applies to his/her included stocks (perhaps zero and one) accounts for any observed excess returns.

These studies are the latest to use the manager's relevant universe as a starting point or benchmark. Grinblatt proposes that successful managers hold securities when their returns are higher than normal and suggests that held securities' returns be compared to the same securities' returns when they are not held. He also suggests the use of a benchmark chosen from the manager's own portfolio by recording the stocks held, the weighting of the stocks and the returns when held. These stocks and weightings can then be used to calculate returns in other periods. This is similar to what we have done in this study.

Grinblatt and Titman empirically examine the issues raised in Roll (1978), Jensen (1972) and Dybvig and Ross (1980) and implement the above strategy by analyzing the performance of 109 "passive portfolios" and 279 mutual funds. The 109 portfolio control group is assumed to have no abnormal performance and is used to test for bias. The authors test several different models in an attempt to determine relationships between returns and various portfolio attributes (size, beta, dividend yield, turnover, past returns). They found that there was a high degree of correlation among the different tests, and that there was a positive relationship between past performance and future performance and between turnover and return. They found no relationship between return and size or between return and expenses. Grinblatt and Titman warn that this study must be viewed with caution due to biases in models and benchmark selection, and suggest that further research of this type be undertaken.

CHAPTER III

HYPOTHESES AND JUSTIFICATION

The hypotheses which are tested in this dissertation are all designed to explore the question of whether or not there is value in the management of equity portfolios. We hope to determine whether or not there is evidence of stock selection skills and information processing ability among mutual fund managers. This study speaks to the market efficiency question and to the value of information. While we speak in the language of investment professionals with such phrases as "beating the market" and earning "excess returns," we are really trying to discover something much more basic and much more important to the market system.

The important question is not the (perhaps selfish) one of knowing whether a given individual or group can earn excess returns through some gaming process, but the question of the efficient allocation of resources. Market economists state that capital will flow to those uses that are most cost effective and most desired by society.

Excess returns are earned by finding those areas, those investments, those firms where capital has been underallocated. In seeking excess returns, managers are merely trying to discover areas in which the consumers' dollar votes are demonstrating or are about to demonstrate their preference for certain products and/or services. It is the skill, the insight, the superior information gathering, collating, analyzing and interpreting ability that would enable managers to earn returns which are above those of the average investment. It is this skill that we are trying to identify and quantify.

Differences in the conclusions drawn by prior studies were due at least in part to the use of external comparison data. The results of any given test have been shown by Roll (1978), and Dybvig and Ross (1980) to be highly dependent on the choice of the comparison data and by Levy (1972) to be highly sensitive to the time period and interval chosen for base period data estimation. The tests devised in this study are straightforward and suffer from none of the deficiencies noted in the prior studies. There will be, of course, some sampling error due to missing observations, but there is no a priori reason to believe that these errors are other than random.

The first set of tests deals solely with stocks bought, sold and held by each fund. Since there is no external benchmark, the problem of appropriateness does not arise. In addition, since these "hold," "buy" and "sell" portfolios are being compared to each other in the same time frame, we can even avoid the problem of return generating process stationarity and time period sensitivity. There is always a possibility that a given fund may have an increasing or a decreasing return over the test period due to steadily increasing or decreasing risk, but this is virtually impossible for the sample as a whole because the sample itself is a reasonable proxy of "the market." The sampling error problem is also extremely small due to the very large sample size used in this study.

The hypotheses are

$H0:_1$ There is no statistical difference between the returns earned on the stocks held by mutual fund(s) at the end of a given quarter and the stocks bought by mutual fund(s) during that same quarter over the next three, six and twelve months.

This hypothesis compares the returns earned on stocks bought by a given fund in a given quarter with returns on stocks held by the same fund during the same quarter. Presumably, value maximizing managers will attempt to buy those stocks (within their relevant universe) which are believed to offer higher returns than other acceptable stocks. There are two cases where a manager might purchase stock in a given time period. (1) A surplus of cash is available due to fund purchases by investors exceeding redemptions. (2) Although no "natural" surplus of cash is available (as in case 1), fund managers sell shares of one stock in order to purchase shares of another.

In case (2), the manager is making a very strong statement that the purchase is expected to perform unusually well, as the manager is willing to spend two transactions costs to make the change. Even in case (1), a manager should buy the stock that is considered to be the best alternative. Such pure "liquidity" trades should still be made with the objective of maximizing the portfolio's return.

$H0{:}_2$ There is no statistical difference between the returns earned on the stocks held by mutual fund(s) during a given quarter and the returns on the stocks sold by mutual fund(s) during that same quarter over the next three, six and twelve months.

As in hypothesis $H0_1$, the returns of two "portfolios" are compared. In this case, the portfolios are (1) all stocks held by a given fund at the end of a given quarter and (2) all stocks sold by the fund during the same quarter. The same reasoning as was used in $H0_1$ also applies here. The manager should sell those issues in which there is the least confidence. The stock may be sold in a given quarter strictly for liquidity purposes such as to pay cash to investors redeeming shares or the stock may be sold even though the cash is not needed immediately. In the former case the manager sells stocks from those owned, choosing the least promising. In the latter case, the manager sells some stock and either holds the cash released or buys other stocks. This latter case would seem to be a stronger indication that the stock sold is out of favor, at least in the eyes of this manager.

Returns on the "sell" portfolio that are significantly lower than those on the "hold" portfolio would provide evidence of manager skill or knowledge (or both).

H0:$_3$ There is no statistical difference between the returns earned on the stocks bought by mutual fund(s) during a given quarter and the returns on stocks sold by mutual fund(s) during that same quarter over the next three, six and twelve months.

This test is like a combination of the above two. Here, stocks bought in a given quarter are compared directly with those sold. In theory, this should be the strongest test. The major weakness is that the numbers of stocks bought and sold are less than those held during the same time period. With the extremely large sample, however, this is not a significant problem. Obviously, when there is neither a shortage nor a surplus of cash, buys and sells taking place simultaneously indicate a strong preference for the stocks purchased. It can be argued that even though the buys and sells occur in the same quarter, they are still executed for liquidity purposes which, in turn, is caused by the fund's purchases and redemptions not being perfectly matched. This is, however, a rather weak argument as funds normally keep a substantial cash reserve. (Five to ten percent is typical for managers who do not engage in market timing, but, instead, only stock selection.)

If this test indicates that stocks bought earn significantly higher returns than stocks sold, it will provide strong evidence for manager skill and/or superior information.

$H0:_4$ There is no statistical difference between the returns earned on the actual stock holdings of mutual funds over 33 quarters (1 January 1977 through 31 March 1985) and the returns over that same time period of a portfolio consisting of the dollar weighted average stock holdings of each manager.

This hypothesis tests actual fund returns over the total test period against a benchmark formed from each fund's holdings over the test period. The index is formed by building a hypothetical portfolio where each component is the average dollar holding of a stock over the test period [$X_i = (\sum_{i=1}^{n}$ Value of stock i in period n$)/33$]. The sum of these individual stock holdings $(\sum_{i=1}^{n} X_i)$ comprises the index. This benchmark follows the suggestions of Cornell (1979), Ferguson (1980), and Roll (1980) that the best benchmark is formed from the evaluated manager's actual holdings with the exception that this test chooses the benchmark from the test period rather that from a prior or later period. A detailed explanation of this index and a numerical example are included in the "methodology" section.

Given that the manager may hold this set of stocks in any positive amount (including zero), can the manager vary the weighting of the stocks to earn returns greater than those that would have been earned by the same stocks held in the index amounts? As in the above hypotheses, we can argue that some trading must take place strictly for liquidity purposes, but this could be done (within very narrow bounds) by varying all holdings proportionately to satisfy redemptions and absorb purchases. Clearly, this is not the way that mutual fund managers operate. It seems reasonable to assume that the majority of trades are made with the intention of improving future performance.

If actual returns earned in this test statistically exceed index returns for a large number of funds (and for the sample as a whole) we would have support for the position that mutual fund managers do possess special skills and/or knowledge.

$H0:_5$ There is no statistical difference between the returns earned on the actual stock holdings of mutual funds over 33 quarters and the returns that would have been earned on each fund's holdings as of the end of the first quarter.

Here we are testing actual returns against another index. This test incorporates a true "buy and hold" strategy. The index in this case is simply each fund's holding at the start of the test period. Changes from this initial holding can be attributed to several different causes. First, since these are open ended funds, changes can be necessitated by purchases not being equal to redemptions in a given time period. Second, some stocks might cease to exist or change their identity during the test period. This test makes no direct correction for this occurrence, but to test for bias, a portfolio is built from the stocks with information available for less than the full 33 quarters and the mean return on the portfolio thus formed is compared to the mean return of all stocks with information available in all test quarters. The difference in mean returns on these portfolios was insignificant.[1]

The last and most important change for this test comes about because fund management's expectations change. Deliberate adjustments are made to the initial holdings due to events both extant and anticipated that increase management's expected returns while maintaining the fund's primary focus (growth, income, etc.). Once again, it is the skill of fund managers in making these deliberate adjustments that we are trying to examine. If the fund's actual returns are significantly greater than index returns, then we once again have evidence of management skill.

All of these hypotheses, then, are designed to test the skills and/or special knowledge of mutual fund managers. Taken together, they should add to the literature examining the purported value of portfolio management. A major difference between this study and previous studies is the data. With two exceptions, Grinblatt and Titman and Copeland and Mayers, no previous studies have used the actual equity securities held by investment managers, and in the studies of Value Line by Copeland and Mayers, only one investment manager was evaluated. The availability of actual equity holdings provides a number of advantages. First, a more accurate measurement of "stock selection" returns is available than from measures such as Jensen's alpha which tends to intermix the impacts of stock selection and market timing. Second, risk measures are of less importance in the first and second test methodologies because each portfolio (mutual fund) is measured against itself rather than against an index. Each fund is, in a sense, its own index. Third, there is much less a problem with non-stationarity of returns in the first methodology than with standard tests such as those of Sharpe, Treynor and Jensen. Finally, the logical simplicity of the approach could be considered a benefit. We test very simple ideas. In particular, 1) Are subsequent returns on buys, sells and holds statistically different? 2) Are the returns on stocks actually held and the dollar weighted average holdings statistically different? 3) Are the returns on stocks actually held and stocks initially held statistically different?

---

[1] The difference was insignificant at the 5% level (T = -0.22).

CHAPTER IV

DATA

Data Sources

The data for this study are taken from several sources. These sources are listed below:

1. Spectrum Tapes. These tapes are issued quarterly and contain the common stock holdings and transactions of all registered investment companies filing NIQ reports with the SEC. All NYSE, ASE and virtually all NASDAQ stocks are included. SPECTRUM (1 & 2) tapes contain four files. The information in these files includes each fund name, fund number, end-of-quarter assets, stock names, cusips, end-of-quarter stock price, number of shares of each stock held at the end of each quarter and shares purchased or sold during the covered quarter. A total of 33 SPECTRUM tapes were utilized in this study (January-March 1977 through January-March 1985). The SPECTRUM TAPES are available from CDA Investment Services in Silver Springs, Maryland. Although not previously used by

academic researchers, they are now available at UCLA and the University of Florida.

2. Price, Dividend, Earnings Tapes. These tapes contain data on stocks listed on the NYSE, ASE and virtually all NASDAQ stocks. The data include monthly price and dividend information on each stock. The "PDE" tape is available from Standard & Poors.

3. CRSP Tapes. As PDE tapes do not include data on banks and insurance companies, CRSP information was used to fill in the price and dividend information for these firms.

4. PDE Update Publications. These (hardcopy) publications contain a listing of firms covered by the PDE tapes including changes, additions and deletions. These publications (issued monthly) were used to keep track of firm name and cusip changes as well as to ferret out merging firms.

5. NYSE, ASE and OTC Daily Price Indexes. These (hardcopy) publications were used to fill in information not contained on any of the tapes and to help in the computation of returns of firms engaged in acquisitions and mergers, especially in the quarter when the change took place.

6. Moody's Handbook of Common Stocks. Moody's was used primarily to determine the dates of splits and mergers as well as to check on other data.

7. The Wall Street Journal Index. This index, along with the articles that were referenced in the Journal, was used to determine the terms of mergers and acquisitions in order to compute returns during the quarters when changes took place.

The Sample

The sample used in this study consisted of all mutual funds that were in existence for the entire study period (1 January, 1977, through 30 June, 1985). Additionally, fund objectives were limited to Income, Growth and Income, Balanced, Growth and Aggressive Growth. Special purpose funds (Bond, Tax Exempt, Gold etc.) were omitted due to the limited return information on the types of securities they own. This set of criteria yielded a generous sample of 293 mutual funds.[1]

Issues Used in the Study

Stocks that were held at any time in all funds.	6485
Stock data contained on PDE tapes.	3827
Stock data contained on CRSP tapes.	1104
Stocks which were unmatched.	1554

---

[1] The selection of funds in this manner subjects the study to sample selection bias (survival bias) which would tend to favor positive results. Grinblatt and Titman, in a paper presented at the 1986 AFA meeting tested for this bias in a similar sample and found none.

For the 1554 unmatched stocks, another search was conducted. Stocks that were 1) held by 10 funds or more, 2) had aggregate holdings of $10 million or more, or 3) a combination of being held by at least 5 funds and having at least $5 million aggregate holdings were the subjects of further search. Two hundred and one stocks met these criteria; 1353 did not and were dropped from the study.

Of the 201 stocks that were unmatched and met the criteria, 45 were missing because of name changes. The data on these firms were gathered and included in the study. Fifty were the subjects of acquisitions and mergers. Data were found on 45 of these, which were also included in the study. The reason that 106 of the 201 firms were unmatched was not discovered, and the data on those firms could not be found. These firms, and the 5 merged firms on which data could not be found, were also dropped from the study. Thus, the vast majority of holdings (5,021 stocks) is included in the study.

CHAPTER V

METHODOLOGY

The research design for this study was constructed to detect investment manager's skill in stock selection.

The study contains three conceptually related yet different approaches. In the first approach, the stock transactions of each mutual fund were formed into three portfolios, and a series of tests were conducted on these portfolios. These portfolios consisted of

(1) All stocks held by each fund at the end of each quarter (1st quarter 1977 through 1st quarter 1985).

(2) All stocks bought during each quarter.

(3) All stocks sold during each quarter.

In the second approach, an index unique to each fund's equity holdings was formed and the returns earned by this index were compared to returns on the actual holdings of the fund. In the third approach, returns that would have been earned if all mutual funds had retained their portfolios held on 31 March, 1977, for the entire test period were compared to actual fund stock holding returns.

The First Approach

The first approach examines returns earned on three portfolios in a pairwise comparison. These portfolios were made up of stocks held at the end of each quarter, as well as stocks bought and sold during each quarter by each fund. The returns for the next three, six and twelve months were calculated on each portfolio and these returns were compared--1) Holdings vs Purchases, 2) Holdings vs. Sales and 3) Purchases vs. Sales.

Returns were calculated as follows:

$$R_3 = (P_{t+1} + D_{t+1})/P_t$$

$$R_6 = (P_{t+2} + D_{t+1} + D_{t+2})/P_t$$

$$R_{12} = (P_{t+4} + D_{t+1} + D_{t+2} + D_{t+3} + D_{t+4})/P_t$$

where

t = the end of a given quarter,

R_3 = the quarterly return on a given stock held at the end of a given quarter,

R_6 = the six month return on a stock held at the end of a given quarter,

R_{12} = the twelve month return on a stock held at the end of a given quarter,

P_t = the price of the stock at the end of the examination quarter,

P_{t+1} = the price of the stock at the end of the quarter following the examination quarter,

P_{t+2} = the price of the stock six months from the end of the examination quarter,

P_{t+4} = the price of the stock twelve months from the end of the examination quarter,

D_{t+1} = the total dividends paid during the quarter following the examination quarter,

D_{t+2} = the total dividends paid during the 2nd quarter following the examination quarter,

D_{t+3} = the total dividends paid during the 3rd quarter following the examination quarter,

D_{t+4} = the total dividends paid during the 4th quarter following the examination quarter.

If we were calculating the return on a stock held on 31 March, 1977, P_{t+1} would be the price of the stocks on 30 June, 1977.

Thus, the returns take capital gains and dividends into account. We calculated the returns on the 5021 stocks (204,000 observations).

Fund equity holding, buy, sell returns were then calculated for each fund as follows: First, the number of shares held (bought, sold) at the end of each quarter was multiplied by the stock price for each stock. This product is the market value of each stock at the end of each quarter. Next, each market value thus calculated is multiplied by the return on that particular stock over the succeeding three, six and twelve months (the R_3, R_6, and R_{12} values above). This product gives the values of each holding if it were retained for three, six and twelve months. This results in beginning and ending values for each stock for each time period in question. The beginning and ending values for each stock were then summed (for each fund, for each quarter) yielding beginning and ending values for all (included) stocks for each time period in question. It was then a simple task to divide the end of period values by the beginning of period values to obtain a return relative for each period for each fund. Table (1), below, illustrates this calculation.

Symbolically

For quarterly returns: $R^t_{3,f} = \sum_{i=1}^{n} P^t_i \; X^t_i \; R^t_{3,i} / \sum_{i=1}^{n} P^t_i \; X^t_i$

For six month returns: $R^t_{6,f} = \sum_{i=1}^{n} P^t_i \; X^t_i \; R^t_{6,i} / \sum_{1=1}^{n} P^t_i \; X^t_i$

For twelve month returns: $R^t_{12,f} = \sum_{i=1}^{n} P^t_i \; X^t_i \; R^t_{12,i} / \sum_{i=1}^{n} P^t_i \; X^t_i$

where X^t_i is the number of shares of stock i held at time t.

TABLE 1: EXAMPLE: FUND XYZ HOLDINGS

QTR 7703

CUSIP	PRICE	#SHARES	R3	R6	R12
12345	10	100	1.1	0.95	1.2
23456	15	200	1.2	1.0	1.1
34567	20	300	1.1	1.2	1.3

MULTIPLY PRICE*SHARES (PRICE*SHARES = VALUE0)

QTR 7703

CUSIP	VALUE0	R3	R6	R12
12345	$ 1,000	1.1	0.95	1.2
23456	3,000	1.2	1.00	1.1
34567	6,000	1.1	1.2	1.3

MULTIPLY VALUES*RETURNS (VALUE0*R3 = VALUE3 etc.)

QTR 7703

CUSIP	VALUE0	VALUE3	VALUE6	VALUE12
12345	$ 1,000	$ 1,100	$ 950	$ 1,200
23456	3,000	3,600	3,000	3,300
34567	6,000	6,600	7,200	7,800

SUM VALUES AT EACH TIME

QTR 7703	SUMV0	SUMV3	SUMV6	SUMV12
	$ 10,000	$ 11,300	$ 11,150	$ 12,300

CALCULATE RETURNS (RTN3 = SUMV3/SUMV0 etc.)

QTR 7703	RTN3	RTN6	RTN12
	1.13	1.115	1.23

With returns calculated for each quarter for each fund, quarterly (6 month, 12 month) returns on the buy, sell and hold portfolios in each quarter were compared by taking the differences between the returns on the buy, sell and hold portfolios in a paired comparison (buy-sell, buy-hold, hold-sell). The means of these differences were then taken for each fund and the differences tested for significance with a Student's T-test. This procedure yielded seven tests for each paired comparison with 293 observations in each test (one quarterly test, two six-month tests, four twelve-month tests). The total number of tests in this section was 7*3 or 21. Two different six-month and four different twelve-month tests were conducted to obtain independent samples. The first six-month test included the six-month returns earned by each fund on equities held (bought, sold) in the first and third quarters of each year. The second six-month test utilized the second and fourth quarters. If all quarters were used in one test, our samples would not have been independent because test elements would be included in more that one observation (for example, the third quarter returns would be included in the first and second observations). The twelve-month tests had to be separated for the same reason. To compare individual fund returns, quarterly (6-month, yearly) returns were compared using a means comparison test (33 observations for three-month returns, 16 or 15 observations for six-month returns and 8 or 7 observations for twelve-month returns). In addition to these aggregate tests, approximately 6,150 tests were also conducted on individual funds in this section.

The Second Approach

The second approach (methodology) compared the returns actually earned by the stocks held in each fund in each quarter to the returns that would have been earned on an index. The impetus for this test was provided by papers by Cornell, by Ferguson and by Roll[1] who took issue with the current practice of comparing portfolio performance to arbitrarily chosen indices. These authors (among others) argued that it is unfair and indeed inappropriate to compare the returns earned by portfolios with different constraints and objectives to a common index or with each other. Instead, these authors suggested that each portfolio or manager be evaluated against an index constructed especially for that manager. Additionally, that index should be built from those instruments in which the manager normally deals--the available set. The method of choosing these unique indices in this study follows the theory of revealed preference; i.e., the available set for each manager is constructed from those assets actually held during the test period. Partly to avoid the criticisms directed at universal indices and the CAPM, but mainly to comform to the suggestions cited above, a unique index was constructed for each portfolio as follows:

---

[1] In "Ambiguity When Performance is Measured by the Securities Market Line" Roll (1978, p.1062), states "No portfolio manager, regardless of his skill, could possibly pick a superior portfolio. Notice, however, that many observed, managed portfolios might also be ex ante efficient and thus be just as good as the index, (gross of fees). This fact could not be determined by using the securities market line criterion with the original index, of course, but it could be determined by employing each managed portfolio as the index in another contest." In "The Expected Return Benchmark and Portfolio Performance Evaluation" Roll (1980, p.22), writes "In order to detect and correct an error in the CAPM benchmark, the portfolio management evaluator must obtain an independent estimate of the error's two components, B and y. This is tantamount to obtaining independent estimates of the evaluated portfolio's expected return. One fairly straightforward method for obtaining such estimates is to apply the classification scheme ... to the individual securities approved for purchase by the portfolio manager. During some validation period (different from the period of management evaluation), each approved security would provide a sample estimate of B and y from observed rates of return. These y and B estimates must be calculated with the market index that will ultimately be used in performance evaluation."

1. Determine the number of shares of each stock held in a fund at the end of each quarter (March 1977 - March 1985).

2. Determine the unadjusted price of each stock held by a fund at the end of each quarter. (The unadjusted price is the price that existed at that point in time unadjusted for subsequent splits and stock dividends.)

3. Multiply the price and share figures together for each stock for each quarter. (We now have the value of

each stock held in each quarter - the V0 value in test one.)

4. Add together the values of each stock held in each quarter for the 33 quarters of the test period to yield the total dollar value of each stock held in the test period, (i.e., $\sum_{t=1}^{33} P_{i,t} X_{i,t}$ for stock i where $P_{i,t}$ is the price per share of stock i in time t, and $X_{i,t}$ is the number of shares of stock i held in time t).

5. To obtain the average dollar value held, divide the total obtained in step four by 33 for each stock.

6. The aggregate of these stock value averages is taken as the index against which performance will be measured.

Symbolically, for each stock: $\sum_{t=1}^{33} P^f_{i,t} X^f_{i,t} /33 = V0^f_i$

where $P^f_{i,t}$ is the price of stock i at the end of quarter t in fund f,

$X^f_{i,t}$ is the number of shares of stock i at the end of quarter t in fund f,

$V0^f_i$ is the index value of stock i in fund f.

The argument supporting this index is that the unbiased estimate of a fund's intended set of potential equity holdings is the set of securities actually held by the fund over time.

To find quarterly index returns, the index values of each stock were added to obtain a total beginning index value.

Beginning of each quarter index value for fund f

$$(VO_f) = \sum_{i=1}^{n} VO^f{}_i$$

This sum is the same for each quarter. If a quarterly return were available for each stock in each quarter, then it would be a simple matter to multiply each stock's value by the return relative for that quarter, obtain an end-of-quarter value for each stock, sum the end-of-quarter values and divide this total by the index value (beginning-of-the-quarter value) to obtain a fund quarterly return.

$$\sum_{i=1}^{n} VO_i R_i / \sum_{i=1}^{n} VO_i$$

Due to occasional missing returns, another approach was used. Quarterly index returns were determined by taking the dollar weighted average return for each quarter utilizing the stocks with available returns. While this procedure is a potential source of error, there is no reason to believe that errors thus introduced should be systematically biased in any particular direction. In addition, missing observation bias was tested for, and found to be insignificant (footnote 1, Chapter IV). These errors, then, should tend toward zero for a large number of observations.

7. The actual returns for each fund were calculated in the same manner as the quarterly tests in methodology one.

8. The index return (adjusted for missing observations) for each quarter was then subtracted from the actual return for the same quarter yielding 33 return differences.

9. The differences obtained in 8 were tested (with a Student's T-test) to determine whether or not they were significantly different from zero.

The results of these tests are listed and explained in the results section. A numerical example of this test is delineated below.

The overall T-test in this instance is conducted by multiplying the quarterly actual/index returns together to obtain 33-quarter return relatives and subtracting the index return relative from the actual return relative for each fund, to obtain 293 differences. Since there are a large number of observations, and since these observations are mound shaped, a T-test is appropriate to test whether or not the mean of these differences is statistically different from zero.

TABLE 2: TEST 2 EXAMPLE: FUND 999

QTR 1

STOCK	PRICE	SHARES	VALUE	RETURN
1	10	100	$1,000	0.9
2	11	200	2,200	0.91
3	12	300	3,600	0.92
4	13	400	5,200	0.92

QTR 2

STOCK	PRICE	SHARES	VALUE	RETURN
1	9	100	$ 900	1.11
2	10	100	1,000	1.20
3	11	200	2,200	1.18
4	12	300	3,600	1.17

QTR 3

STOCK	PRICE	SHARES	VALUE	RETURN
1	10	100	$1,000	1.10
2	12	100	1,200	0.83
3	13	200	2,600	0.92
4	14	400	5,600	0.86

QTR 4

STOCK	PRICE	SHARES	VALUE	RETURN
1	11	100	$1,100	1.06
2	10	0	0	1.10
3	12	300	3,600	1.10
4	12	200	2,400	1.14

TOTAL VALUES: (VTOTAL = VQTR1 + VQTR2 + VQTR3 + VQTR4)

STOCK	VQTR1	VQTR2	VQTR3	VQTR4	VTOTAL
1	$1,000	$ 900	$1,000	$1,100	$ 4,000
2	2,200	1,000	1,200	0	4,400
3	3,600	2,200	2,600	3,600	12,000
4	5,200	3,600	5,600	2,400	16,800

TABLE 2 (Continued)

AVERAGE VALUE OF EACH STOCK HELD (VTOTAL/4)

STOCK	VTOTAL	VAVERAGE
1	$ 4,000	$1,000
2	4,400	1,100
3	12,000	3,000
4	16,800	4,200

TABLE 3: FUND 999 RETURN IF ALL STOCKS HAD RETURN INFORMATION AVAILABLE IN ALL QUARTERS

QTR1

STOCK	RETURN	AVE VALUE	END OF QUARTER VALUE
1	.9	$1,000	$ 900
2	.91	1,100	1,000
3	.92	3,000	2,760
4	.92	4,200	3,860
SUM		$9,300	$8,520

QUARTER 1 RETURN RELATIVE = $8,520/$9,300 = 0.916

QTR2

STOCK	RETURN	AVE VALUE	END OF QUARTER VALUE
1	1.11	$1,000	$ 1,110
2	1.20	1,100	1,320
3	1.18	3,000	3,540
4	1.17	4,200	4,914
SUM		$9,300	$10,884

QUARTER 2 RETURN RELATIVE = $10,884/$9,300 = 1.17

Table 3 (Continued)

QTR3

STOCK	RETURN	AVE VALUE	END OF QUARTER VALUE
1	1.10	$1,000	$ 1,100
2	0.83	1,100	913
3	0.92	3,000	2,760
4	0.86	4,200	3,612
SUM		$9,300	$ 8,385

QUARTER 3 RETURN RELATIVE = $8,385/$9,300 = 0.9

QTR4

STOCK	RETURN	AVE VALUE	END OF QUARTER VALUE
1	1.06	$1,000	$ 1,060
2	1.10	1,100	1,210
3	1.10	3,000	3,300
4	1.14	4,200	4,788
SUM		$9,300	$10,358

QUARTER 4 RETURN RELATIVE = $10,358/$9,300 = 1.114

With these return relatives, we developed a test period geometric index return for each fund by multiplying the return relatives together. For this example, fund 999, the index return is (.916 * 1.17 * .9 * 1.114) = 1.075.

We also determined a geometric average return by extracting the N^{th} root (in this case N = 4). Fund 999's geometric average return relative is $(1.075)^{1/4}$ = 1.018 or a 1.8% rate of return.

Unfortunately, returns were not available for each stock in every quarter, necessitating an adjustment of the available data to make it comparable to the actual returns earned by each fund during the test period. The adjustment would change the example as follows: Let us assume that the return for stock 2 in quarter 4 is missing. This missing return would bias the index downward if no correction were made.

TABLE 4: RETURN CALCULATION WITHOUT ADJUSTMENT

QTR4

STOCK	PRICE	SHARES	VALUE	AVERAGE VALUE	END OF QTR VALUE	RETURN
1	$11	100	$1,100	$1,000	$1,060	1.06
2	10	0	0	1,100	MISSING	MISSING
3	12	300	3,600	3,000	3,300	1.10
4	12	200	2,400	4,200	4,788	1.14
SUM				$9,300	$9,148	

QUARTER 4 RETURN RELATIVE = $9,148/$9,300 = 0.98

INDEX RETURN RELATIVE = .916 * 1.17 * .9 * .98 = 0.945

It is obvious that some sort of adjustment is needed. The adjustment chosen simply allows for the missing data by assuming that the dollar value of the missing stock is divided proportionately among the remaining stocks (the ones that do have data available). The two principal explanations for missing data are (1) the stock was not in existance due to its being a new firm at some time after the test period started, and (2) the firm had gone out of existance due to merger, failure, etc. There is no a-priori reason to suspect systematic bias in the first case, and in the second case one can make the reasonable assumption that the proceeds from the discontinued firm are invested proportionately in the remaining index stocks. Applied to this example, the adjustment is made as follows.

TABLE 5: RETURN CALCULATION WITH ADJUSTMENT

QTR4

STOCK	PRICE	SHARES	VALUE	AVERAGE VALUE	END OF QTR VALUE	RETURN
1	$11	100	$1,100	$1,000	$1,060	1.06
2	10	0	0	1,100	MISSING	MISSING
3	12	300	3,600	3,000	3,300	1.10
4	12	200	2,400	4,200	4,788	1.14
SUM				$9,300	$9,148	

The quarter 4 weighted average return relative is now

(1000/8200)(1.06)+(3000/8200)(1.1)+(4200/8200)(1.14)=1.116

The fund 999 index return is now .916*1.17*.9*1.116=1.076

The geometric average return is $(1.076)^{1/4} = 1.018$.

The Third Approach

A third test was conducted to answer the question: How do returns actually obtained (on equities actually held) compare to the return that would have been earned if each fund had retained its initial set of holdings for the entire period?

This test was done much in the same manner as was the second test as this, too, is a kind of index comparison. As in the second test, geometric index and actual returns were calculated and compared. This test was done in the following manner.

1. The initial holdings (stocks held on 31 March, 1977) for each fund were determined and the dollar value of those holdings were calculated.

$$VO_i = (PO_i)(X_i)$$

TABLE 6: FUND XYZ INITIAL HOLDING

STOCKS	SHARES	PRICE
A	100	$10
B	200	18
C	300	12

Multiply shares*price to obtain initial value:

TABLE 7: INITIAL VALUE

STOCK	VALUE0
A	$1,000
B	3,600
C	3,600

2. The returns on each stock held by a given fund on March 31, 1977, were calculated for each quarter in the test period (a total of 33 quarters).

TABLE 8: FUND XYZ STOCK PRICES AND DIVIDENDS

STOCK	P7703	P7706	D7706
A	\$10	\$11	\$1
B	18	20	1
C	12	11	1

where

P7703 is the price of the stock on March 31, 1977;
P7706 is the price of the stock on June 30, 1977;
D7706 is the sum of all dividends paid after March 31, 1977 and on or before June 30, 1977.

The return on stock A for quarter one is calculated as follows.

R_1 = (P7706-P7703+D7706)/P7703 = (11-10+1)/10 = 1.2 or 20%

The return relatives for each stock were then multiplied together to determine a return relative for the entire test period (March 31, 1977 through March 31, 1985). $(R_i = \sum_{t=1}^{33} R_{i,t})$

4. The final return relative for each stock was multiplied by that stock's initial value to determine an ending value for that stock. (This gives us the value that would have obtained had the stock been held for the entire test period.) $(VO_i \; R_i)$

5. The initial and final values for all stocks in a given fund were summed. $(\sum_{i=1}^{n} VO_i \; R_i$ and $\sum_{i=1}^{n} VO_i)$

TABLE 9: FUND XYZ VALUE

STOCK	VALUE0	RETURN	VALUEF
A	$1,000	5.126	$ 5,126.00
B	3,600	6.034	21,727.40
C	3,600	4.232	15,235.20
SUM	8,200		$41,596.60

6. A return relative can then be calculated for each fund by dividing the fund final value by the initial value for each fund

$$\left(\sum_{i=1}^{n} VO_i \, R_i \, / \sum_{i=1}^{n} VO_i \right)$$

FUND RETURN = VF/V0 = $41,596.60/$8200.00 = 5.69

and a geometric average quarterly return may be calculated for each fund.

AVERAGE RETURN = $(\text{FUND RETURN})^{1/n} = (5.069)^{1/33} = 1.0504$

or a quarterly return of 5.04%.

This test has a source of possible bias in that all holdings did not have subsequent returns for each of the test (33) quarters. Various adjustments were considered, including the one used in test two, but so few stocks (9) actually had missing data that it was decided to use only those stocks for which the data was complete.

To test for bias, those stocks with missing data were formed into a portfolio and the mean return on this portfolio (during the times when data was available) was compared to the mean return of the stocks with complete data. The difference between these means was found to be insignificant, suggesting that this source of bias is minimal.[2] It should be noted that any adjustment to include the stocks with missing data would itself introduce a (perhaps unknown) bias in the results as assumptions would have to be made concerning the periods with missing returns.

As in the first series of tests, the tests of H_4 and H_5 were repeated for each fund. For H_4, the actual returns for each quarter were compared to the index returns for corresponding quarters using a Student's T-test. The results of these tests (33 observations) are presented in the "Results" section. The results of the individual fund tests of H_5 are available from the author on request.

Other Tests

Several other tests were conducted in an attempt to discern relationships between and among the several tests. These tests included Sharpe, Treynor and Jensen tests, comparisons of significance by fund stated objective and tests of consistency or predictability.

---

[2] The difference between means was 0.0309. The standard error was 0.0364, so that the difference was 0.85 standard errors away from the mean.

Sharpe, Treynor and Jensen Tests

Sharpe, Treynor and Jensen tests were run on all funds in the traditional manner. Fund returns in excess of the risk free rate (proxied by 90 day T-bills) were regressed against the excess returns on the S&P 500 over the risk free rate (33 quarterly observations). Correlations were run on the results of these tests using Pearson's product moment correlation. Correlations were also run on several other tests using the mean quarterly (six-month, twelve-month) return differences on the buy-sell, buy-hold and sell-hold tests along with the Sharpe, Treynor and Jensen tests. These correlations included the data on all 293 funds.

Tests by Stated Fund Objective

Mutual funds are commonly classified both by the funds themselves and by outside parties according to their stated objectives. These objectives are meant to describe the fund management's risk tolerance and therefore the fund's expected return. The funds in this study were broken down into groups by objective to determine whether or not funds with different objectives (a reasonable proxy for risk) had different probabilities of achieving significance. Objectives tested were Aggressive Growth (15 funds), Growth (127 funds), Growth/Income (67 funds) and Income (68 funds). The numbers of significant funds in the quarterly tests (buy-sell, buy-hold, sell-hold) were compared to ramdom expectations.

If fund objective had no relationship to significance, we would expect to find 43.3% of funds to be Growth funds, 5.1% to be Aggressive Growth, 22.9% to be Growth/Income and 23.2% Income (5.5% fall into other catagories). In each sample, the actual distribution of significant funds was compared to the expected distribution as follows.

Calculate the standard error $\sigma_{Ps} = (P_h \, Q_h \, /n)^{1/2}$

and the Z statistic $Z = (P_s - P_h)/ \sigma_{Ps}$

where

σ_{Ps} is the standard error of the sample;

P_h is the expected probability;

Q_h is $1 - P_h$;

n is the total number of significant funds;

P_s is the actual observed decimal of significant funds in the catagory being tested.

Example:
Total number of funds . . . 100
Number of Growth funds . . 50
Number of significant funds . 10

Expected number of significant Growth funds 5
Expected probability that a significant fund will be a growth fund 5
Actual number of significant Growth funds in sample 8

Calculation of test statistic: $\sigma_{Ps} = [(.5)(.5)/10] = 0.158$

$$Z = (.8 - .5)/\ 0.158 = 1.9$$

Test of Consistency

To test for consistency, the buy-sell quarterly data was used and the test period (March, 1977--June, 1985) was divided into two time periods. To obtain periods of equal length, the 33rd observation was dropped leaving two four-year (16 quarter) time periods. The objective of the test was to predict the second period results using the results of the first period tests. Two separate, but related, tests of consistency were run. In the first, the differences between the mean return on stocks sold and stocks bought were ranked from 1 to 293 (by fund). In addition, the actual differences were recorded in rank order and the T-statistics were ranked in the same manner. The relationship between the two periods was tested with Spearman's rank order correlation. The second test placed the funds into deciles according to their first period return differences (buy-sell). The second period decile placement was also recorded and the fund placements in the two periods were compared.

CHAPTER VI

RESULTS

Introduction and Summary

Results of Tests of Hypotheses

This dissertation was undertaken in an attempt to shed light on the question, "Do portfolio managers demonstrate the ability to differentiate between stocks that will earn different rates of returns over various time periods?" Several hypotheses were formulated to test this question. They are as follows:

$H0:_1$ There is no statistical difference between the returns earned on the stocks held by (a) mutual fund (s) at the end of a given quarter and the stocks bought by (a) mutual fund (s) during that same quarter over the next three, six and twelve months.

$H0:_2$ There is no statistical difference between the returns earned on the stocks held by (a) mutual fund (s) at the end of a given quarter and the stocks sold by (a) mutual fund (s) during that same quarter over the next three, six and twelve months.

$H0:_3$ There is no statistical difference between the returns earned on the stocks sold by (a) mutual fund (s) during a given quarter and the stocks bought by (a) mutual fund (s) during that same quarter over the next three, six and twelve months.

$H0:_4$ There is no statistical difference between the returns earned on the actual stock holdings of mutual funds over 33 quarters and the returns over that same time period of a portfolio consisting of the dollar weighted average stock holdings of each manager.

$H0:_5$ There is no statistical difference between the returns earned on actual stock holdings over 33 quarters and the returns that would have been earned on each fund's initial holdings as of the end of the first quarter.

Hypotheses $H0:_2$, $H0:_3$ and $H0:_5$ were rejected at the 5% level, lending support to the position that managers have some skill or knowledge. $H0:_1$ was rejected by four of the seven tests. Three of the twelve-month tests failed to reject, but these are the weakest of the tests due to fewer observations in the individual tests and higher standard errors in the overall tests. The results pertaining to $H0:_4$ indicated that the index significantly outperformed the managers. (A separate hypothesis explaining the results of the test of actual results against each manager's index was formulated and tested, explaining the reason for the negative relationship.) The tests of the hypotheses were performed using aggregate data on all managers. In addition, many tests were also performed on each mutual fund.

Discussion

One strategy of this dissertation was to obtain a sample size large enough to uncover differences, if any existed. This strategy appears to have been successful. The three major and two minor data sets included in excess of 750,000 observations of unique fund, quarter, stock combinations, more than enough to operationalize this strategy.

The results of this series of tests were interesting in several ways. In the first place, the outcome of the first set of tests support the same conclusions as the studies cited above by Copeland and Mayers and Black which suggested that managers had the ability to select stocks that would earn returns below their normal returns in the future, but not to select stocks that would earn unusually high returns in the same time frame. Their Value Line studies showed that there appeared to be a difference in subsequent returns among stocks which were rated higher or lower according to Value Line's rating system. Further, the predominant skill in this stock selection process seemed to be the ability to select stocks that would do worse than average rather than better. Next, the second and third tests agreed with studies conducted by Fama, Kon, and Kon and Jen which suggested that managers did possess some stock selection ability, but were unable to time aggregate market movements well. Finally, some new insights were discovered which appear to cast added light on all of these studies. In particular, this study seems to have picked up evidence of a weak ability to select superior stocks that was not obvious in the Value Line studies. While the buy-hold results are the weakest, they are never-the-less present. Also, as a byproduct of the second approach (the index test), we have empirical evidence of the often asserted opinion that managers seem to buy stocks that have been performing well in the past.

Test Methodology One

Results of Buy-Hold-Sell Tests

In each series of tests, the author conducted seven tests: one quarterly, two six-month and four twelve-month tests. The quarterly tests compare the means of the returns on stocks bought or sold during the prior three months or on stocks held at the end of each quarter over the next quarter. The two six-month tests also compare stocks bought or sold during the prior quarter (or stocks held at the end of each quarter), but now the returns are compared over the next six months. To obtain independent samples, the six-month tests starting on 31 March, 1977, used every other quarter (the 1st and 3rd quarters of each year). If every quarter were used, the sample size would have been larger, but the observations would not have been independent. For example, the second quarter data would appear in two sample elements, as would every other quarter except the first and the last. The four twelve-month tests followed the same proceedure as the other tests, and of course four tests were run to obtain independent samples.

Table ten presents the results of test one in this study where stocks bought, sold and held were tested in a paired comparison (buys vs. sales, buys vs. holdings, holdings vs. sales). The test uses all purchases, sales and holdings of all managers. The mean values shown represent 1.0 plus the return. For example, the 1.04477 mean buy return during the next quarter represents a 4.477% quarterly rate of return.

Notice that the way the tests were conducted (buys - sells in each quarter e.g.) results in the mean difference not necessarily being equal to the mean buy minus the mean sell. This is primarily due to missing observations (perhaps there is a quarter in which there are buys, but

TABLE 10: RESULTS OF OVERALL DIFFERENCE BETWEEN MEANS TESTS

STOCKS BOUGHT VS STOCKS SOLD

TEST	MEAN OF BUYS	MEAN OF SALES	MEAN OF DIF (B-S)	T	LEVEL OF SIGNIF
Quarterly	1.04477	1.03590	.00951	5.44	.0001*
6 Mos (7703)	1.08510	1.07259	.01316	4.32	.0001*
6 Mos (7706)	1.09174	1.07584	.01695	4.91	.0001*
12 Mos (7703)	1.17703	1.15597	.02250	3.17	.0017*
12 Mos (7706)	1.19851	1.17846	.01579	1.84	.0668
12 Mos (7709)	1.18630	1.16700	.01804	2.63	.0089*
12 Mos (7712)	1.18679	1.16175	.02881	3.71	.0002*

STOCK HOLDINGS VS STOCKS SOLD

TEST	MEAN OF HOLDINGS	MEAN OF SALES	MEAN OF DIF (H-S)	T	LEVEL OF SIGNIF
Quarterly	1.04370	1.03590	.00715	6.54	.0001*
6 Mos (7703)	1.08762	1.07259	.01024	5.03	.0001*
6 Mos (7706)	1.08874	1.07584	.01083	4.74	.0001*
12 Mos (7703)	1.18229	1.15597	.01812	3.85	.0001*
12 Mos (7706)	1.19079	1.17846	.01101	1.97	.0502
12 Mos (7709)	1.19309	1.16700	.02311	4.82	.0001*
12 Mos (7712)	1.18886	1.16175	.02213	4.15	.0001*

STOCKS BOUGHT VS STOCK HOLDINGS

TEST	MEAN OF BUYS	MEAN OF HOLDINGS	MEAN OF DIF (B-H)	T	LEVEL OF SIGNIF
Quarterly	1.04477	1.04257	.00342	2.60	.0099*
6 Mos (7703)	1.08510	1.08654	.00280	2.01	.0452*
6 Mos (7706)	1.09174	1.08687	.00705	2.36	.0189*
12 Mos (7703)	1.17703	1.17975	.01074	2.02	.0446*
12 Mos (7706)	1.19851	1.18868	.00821	1.43	.1551
12 Mos (7709)	1.17361	1.17686	-.00137	-0.25	.8037
12 Mos (7712)	1.16839	1.67588	.00983	1.59	.1130

* SIGNIFICANT AT THE 5% LEVEL

no sales). Note also that a difference of .00951 translates into a 3.86% annual return difference $[(1.00951)^4]$.

Each comparison consisted of a series of seven tests, one three-month test, two six-month tests and four twelve-month tests. All tests compared the mean returns on stocks bought or sold during each quarter or stocks held at the end of each quarter. The quarterly tests compared the returns over the next three months, the six-month tests over the next six months and the twelve-month tests over the next year. Note that four tests of buys vs. holdings and all but one test of buys vs. sales and all but one test of sales vs. holdings were positive and significant.[1] (Please refer to the statistical section for a thorough discussion of the statistical tests.)

Discussion of Buy-Hold-Sell Test Results

The results of this series of tests give a strong indication that this set of mutual fund managers as a group sold stocks that would return significantly less over the next three-, six- and twelve-month periods than the stocks they retained or purchased. Note also that there is fairly strong evidence that they also purchased stocks that returned more than the stocks they held.[2]

---

[1] The 12-month tests starting in June 1977 had significance levels only slightly above 5% (5.02% and 6.68%) and of course, due to the limited number of observations, the 12-month tests were the weakest.

These results agree with ex-ante expectations as the researcher expected that the decisions would have a hierarchal arrangement, with buys reaping the greatest returns, holds next, and sales returning the least. It seemed that a buy would be made in response to strong positive information , that stocks would be held on more or less neutral information and that stocks would be sold as a result of strong negative signals. In the light of these results, this appears to have been the case, although as did Copeland & Mayers, we find the evidence to be stronger for skill in detecting poor performers.

Results of Individual Fund Tests

In addition to the aggregate tests, all of the tests were repeated for each fund (a total of 6153 tests in this segment). Table 11 summarizes the results of the individual fund tests which are available from the author on request.

---

[2] This series of tests (both the overall tests and the individual fund tests) was also run by finding the mean of the returns on stocks bought-sold-held over the 33 quarters and testing the difference between means for significance. The results of this methodology was essentially identical to the results reported above. The only major difference was that the twelve month buy-sell test starting in June was positive and significant rather than marginally insignificant.

TABLE 11: SUMMARY OF INDIVIDUAL FUND TESTS

TEST	# FUNDS W/+ SIG	% FUNDS W/+ SIG	# FUNDS W/- SIG	% FUNDS W/- SIG
QUARTERLY:				
BUY/SELL	28	9.6	3	1.0
HOLD/SELL	29	9.9	3	1.0
BUY/HOLD	13	4.4	5	1.8
SIX MONTHS:				
BUY/SELL1	15	5.1	3	1.0
BUY/SELL2	20	6.8	3	1.0
HOLD/SELL1	17	5.8	1	0.3
HOLD/SELL2	18	6.1	5	1.7
BUY/HOLD1	8	2.7	7	2.3
BUY/HOLD2	9	3.1	8	2.7
TWELVE MONTHS:				
BUY/SELL1	6	2.1	1	0.3
BUY/SELL2	8	2.7	3	1.0
BUY/SELL3	9	3.1	5	1.7
BUY/SELL4	15	5.2	6	2.1
HOLD/SELL1	16	5.5	2	0.7
HOLD/SELL2	12	4.1	5	1.7
HOLD/SELL3	12	4.1	3	1.0
HOLD/SELL4	12	4.1	1	0.3
BUY/HOLD1	7	2.1	5	1.7
BUY/HOLD2	2	0.9	6	2.1
BUY/HOLD3	8	2.7	9	3.1
BUY/HOLD4	9	3.1	11	3.7

Notes:

1. Six-month tests for stocks bought/sold in the first and third quarters (or held at the end of these quarters) are designated 1, second and fourth quarter tests are designated 2.

2. Twelve-month tests for stocks bought/sold in the first quarter (or held at the end of this quarter) are designated 1, second quarter tests are designated 2, third quarter tests are designated 3, and fourth quarter tests are designated 4.

The level of significance chosen for these tests was 5%, and instances where stocks bought outperformed stocks sold, where stocks bought outperformed stocks held, and where stocks held outperformed stocks sold were deemed positive. (See statistics section.)

The individual fund tests seem to tell the same story as the overall tests, although not as strongly. In every instance of buy-sell and hold-sell tests there are many more funds which are positive and significant than negative and significant (for instance, in the test of quarterly buys vs sells, 28 are positive and only 3 are negative). The buy-hold tests do not exhibit this consistency. The quarterly tests (which are the strongest) have the greatest differences, with about 10% of both the buy-sell and hold-sell tests being positive and significant.

Test for Differences in Risk

Since there is no risk measure embodied in these tests, it is obvious that we are assuming that the risks of the stocks that managers buy are not significantly different from the stocks that they sell over time. While this seems quite reasonable, as systematic differences would result in a fund which would continuously become more or less risky, we decided to test independently for differences in risk. The standard deviation of the quarterly returns on all stocks purchased, sold and held during the test period was chosen as a proxy for risk. The differences in the means of the standard deviations of the quarterly returns of portfolios made up of all stocks bought, held or sold were tested in a paired comparison.

Table 12: RESULTS OF TEST FOR CONSISTENCY AMONG STANDARD DEVIATIONS

TEST	BUYSTD	SELLSTD	HOLDSTD	DIFF	T	PRT
BUY-SELL	0.1232	0.1185	0.0983	0.0047	0.145	94%
BUY-HOLD	0.1232	0.1185	0.0983	0.0249	0.876	69%
HOLD-SELL	0.1232	0.1185	0.0983	-0.0202	-0.750	72%

These tests were done using the quarterly returns of all buys-sells-holds in each fund. The standard deviation of the returns on each of these portfolios for each fund was an element in each observation. Thus there were 293 observations, one for each fund (in each test).

It is obvious from the examination of table (12) that each result was insignificant, tending to support the assumption that there is no difference in the risk of stocks bought, sold or held by the mutual fund managers in this study.

Test Methodology Two

Discussion

The second series of tests compared actual returns to the returns earned on an index unique to each fund (detailed earlier). This set of tests also is consistent with earlier results. If a manager's actual returns significantly exceeded the returns on his particular index (his average holdings) it would indicate an ability to vary the weighting of issues and timing of stock purchases and sales to achieve a greater return than would a buy and hold strategy using the same issues. This is somewhat similar to the skill that was examined by Fama (1972), Kon and Jen (1983) and Kon (1983) whose studies indicated that managers have no (or slightly negative) timing ability. The results of these tests were again consistent with earlier studies in that no evidence of superior stock purchase timing was uncovered.

Results of Test Methodology Two

The results of the overall test of fund actual holdings vs each fund's unique index and a summation of individual tests are presented below in table (13). The results pertaining to each fund are available from the author on request.

It is obvious that the number of funds that underperformed their particular indexes was much greater than those that outperformed their indexes. The majority of fund's results were insignificant.

TABLE 13: RESULTS OF TESTS COMPARING ACTUAL FUND STOCK HOLDING RETURNS TO THE RETURNS ON EACH FUND'S UNIQUE INDEX

Overall T-Test of the Difference in Means Between the Actual Fund Stock Holding Returns and Each Fund's Index Return

Mean of Actual Geometric Return Relative	3.77956
Mean of Index Geometric Return Relative	4.68743
Mean Difference Between Means	-0.90787*
Standard Error of the Mean (Over Funds)	0.14356
T-Statistic	-6.320
Significance Level	0.0001
Number of Funds which are Positive and Significant	12
Number of Funds which are Negative and Significant	109

* Index exceeded actual

Discussion of Test Results

There is a potential explanation for the inordinate number of funds that underperformed their indexes, when in Test Methodology One we found evidence of purchased stocks outperforming sold stocks. If fund managers systematically buy stocks after the stocks have had unusual gains, the fund indexes will tend to show larger returns than the actual portfolios. For example, a stock bought in June, 1980 that had a very unusual return from June to December, 1979 would have the unusual returns included in the fund's index although the stock was in fact not held during the period of high returns.

Lagged Return Test

To test for this systematic bias, returns on stocks purchased were compared to those on stocks held three, six and twelve months prior to the quarter in which they were held or purchased. These lagged return tests supported the hypothesis that managers do indeed systematically purchase stocks after the stocks have had unusual positive returns. The results of these lagged tests are presented in table (14).

TABLE 14: RESULTS OF LAGGED PURCHASES VS HOLDINGS TESTS

TEST	MEAN OF BUYS	MEAN OF HOLDS	BUY MEAN MINUS HOLD MEAN	T	PRT
QUARTER	1.06772	1.06246	0.00526	3.18	0.0016
SIX MONTHS	1.13744	1.13131	0.00613	1.85	0.0660
TWELVE MONTHS	1.31177	1.30017	0.01160	1.57	0.1170

Note that the quarterly test was significant at the one percent level (the annualized difference is 1.005257 - 1= 2.12%). The six- and twelve-month tests were positive, but not significant. The quarterly test compared three-month returns on stocks bought and held three months prior to the test period while the six- and twelve-month tests compared six- and twelve-month returns on stocks purchased six and twelve months prior to the test period. Therefore, the six- and twelve-month differences, while greater than the three-month differences in absolute terms are smaller in relative terms resulting in larger standard errors and smaller T-statistics for these tests.

Test Methodology Three

Discussion

The final test in this study compared actual returns to the returns that would have been earned on each fund's initial holding if the fund had held the stocks that were in the fund as of 31 March, 1977. This test is based on the premise that managers should trade only if they believe the trade (buy or sell) will improve future returns over a no trade strategy. (Liquidity trades could be enacted by buying and selling a proportional amount of the initial holding.) This test is then a true buy and hold comparison. The results of this test are shown in table (14).

TABLE 14: COMPARISON AND TEST OF ACTUAL FUND STOCK HOLDING RETURN WITH THE RETURNS THAT WOULD HAVE BEEN EARNED ON INITIAL HOLDINGS

OVERALL TEST OF DIFFERENCE IN MEANS

ACTUAL RETURN	INITIAL HOLDING RETURN	DIFFERENCE	STANDARD ERROR	T	LEVEL OF SIGNIFICANCE
3.78542	3.56976	0.21566	0.07113	3.03	0.0027

These results represent the total period geometric returns. The quarterly difference between means is 0.005933 (.5933%).

Test Three Results

The results of this test agree with those obtained in the first series of tests. The actual holdings earned statistically greater returns than the "buy/hold" strategy. Note the difference between this test and the index test. The index test would be affected by the tendency of managers to buy stocks after they have made unusual gains, but the initial holding test would not be.

Additional Tests

While they were not a major component of this study, Sharpe, Treynor and Jensen tests were run on each fund. The results of these tests are available from the author. As in previous studies the results of the three tests are quite consistent. With rare exceptions, a fund out-performing another using one test achieved the same results using another test. A correlation matrix of the 293 funds is presented in table (15) below.

TABLE 15: CORRELATION MATRIX OF SHARPE, TREYNOR AND JENSEN MEASURES

	SHARPE	TREYNOR	JENSEN
SHARPE	1.0000 (0.0000)		
TREYNOR	0.4917 (0.0001)	1.0000 (0.0000)	
JENSEN	0.9278 (0.0001)	0.5302 (0.0001)	1.0000 (0.0000)

Note: Numbers in parentheses are the probabilities of this correlation if in fact the correlation were zero.

Summary of Tests

To bring together all results in one place, a matrix of all 293 funds and a sample of all tests is presented in table (16).

Note the consistency among the different sets of tests. While R^2 values were not particularly high [with the exception of the Sharpe-Jensen test (.861)], some interesting relationships were apparent. In each case, the correlations between the Buy-Sell and Hold-Sell tests were stronger than other combinations. Also, in each case, the R^2 values were higher between tests with equal time periods than between tests with different time periods. Comparisons between tests with different time periods favored the tests that were in the same catagories (Buy-Sell quarterly vs. Buy-Sell six-months). This is not surprising, because of some co-linearity between these tests. It is interesting as well to note that the R^2 values are higher in the shorter time periods just as the test T-statistics were higher for these periods. Note that although the Sharpe, Treynor and Jensen tests were all highly correlated with each other, they were not, in general, highly correlated with the other tests.

TABLE 16: CORRELATION MATRIX OF FOURTEEN TESTS ON ALL FUNDS

	BSQ	HSQ	BHQ	BS6	HS6	BH6
SHAR	-0.0021	0.1235	-0.2198	0.0155	0.0809	-0.1699
	(0.9719)	(0.0346)	(0.0001)	(0.7911)	(0.1673)	(0.0035)
TREY	-0.1057	-0.0579	-0.2409	-0.0260	-0.0633	-0.1464
	(0.0709)	(0.3230)	(0.0709)	(0.6573)	(0.2803)	(0.0121)
JEN	-0.0574	0.0727	-0.2798	0.0561	0.1039	-0.1872
	(0.3276)	(0.2149)	(0.0001)	(0.3386)	(0.0759)	(0.0013)
BSQ	1.0000	0.7669	0.5560	0.4809	0.4385	0.0993
	(0.0000)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0890)
HSQ	0.7669	1.0000	0.1742	0.4695	0.6395	0.0560
	(0.0001)	(0.0000)	(0.0028)	(0.0001)	(0.0044)	(0.3391)
BHQ	0.5560	0.1742	1.0000	0.2564	0.0373	0.5177
	(0.0001)	(0.0028)	(0.0000)	(0.0001)	(0.5245)	(0.0001)
BS6	0.4809	0.4695	0.2564	1.0000	0.7252	0.4689
	(0.0001)	(0.0001)	(0.0001)	(0.0000)	(0.0001)	(0.0001)
HS6	0.4385	0.6395	0.0373	0.7252	1.0000	0.0339
	(0.0001)	(0.0001)	(0.5245)	(0.0001)	(0.0000)	(0.5635)
BH6	0.0993	0.0560	0.5177	0.4689	0.0339	1.0000
	(0.0898)	(0.3391)	(0.0001)	(0.0001)	(0.5635)	(0.0000)
BS12	0.1361	0.1991	0.0261	0.1875	0.1359	0.0504
	(0.0198)	(0.0006)	(0.6563)	(0.0013)	(0.0200)	(0.3903)
HS12	0.2591	0.3518	0.1064	0.3758	0.5120	0.0154
	(0.0001)	(0.0001)	(0.0689)	(0.0001)	(0.0001)	(0.7936)
BH12	0.0733	-0.0270	0.2068	0.3048	-0.0482	0.4646
	(0.2109)	(0.6448	(0.0004)	(0.0001)	(0.4115)	(0.0001)
A/IX	-0.1116	-0.0983	-0.0288	-0.1580	-0.1906	0.0304
	(0.0563)	(0.0929)	(0.6237)	(0.0067)	(0.0010)	(0.6048)
A/IN	0.0949	0.1548	0.0820	0.1817	0.2198	0.1090
	(0.1051)	(0.0079)	(0.1616)	(0.0018)	(0.0001)	(0.0624)

TABLE 16: (Continued)

	BS12	HS12	BH12	A/IX	A/IN
SHARPE	0.1741 (0.0028)	0.0482 (0.4116)	-0.0379 (0.5178)	-0.0924 (0.1143)	0.4328 (0.0001)
TREYNOR	0.0981 (0.0936)	-0.0668 (0.2543)	0.0548 (0.3498)	-0.0702 (0.2311)	0.0606 (0.3012)
JENSEN	0.2440 (0.0001)	0.0718 (0.2206)	0.0248 (0.6728)	-0.0909 (0.1206)	0.4375 (0.0001)
BSQ	0.1361 (0.0198)	0.2591 (0.0001)	0.0733 (0.2109)	-0.1116 (0.0563)	0.0949 (0.1051)
HSQ	0.1991 (0.0006)	0.3518 (0.0001)	-0.0270 (0.6448)	-0.0983 (0.0929)	0.1548 (0.0079)
BHQ	0.0261 (0.6563)	0.1064 (0.0689)	0.2068 (0.0004)	-0.0288 (0.6237)	0.0820 (0.1616)
BS6	0.1875 (0.0013)	0.3758 (0.0001)	0.3048 (0.0001)	-0.1580 (0.0067)	0.1817 (0.0018)
HS6	0.1359 (0.0200)	0.5120 (0.0001)	-0.0482 (0.4115)	-0.1906 (0.0010)	0.2198 (0.0001)
BH6	0.0504 (0.3903)	0.0154 (0.7936)	0.4646 (0.0001)	0.0304 (0.6048)	0.1090 (0.0624)
BS12	1.0000 (0.0000)	0.1951 (0.0008)	0.0800 (0.1723)	-0.0578 (0.3244)	0.1117 0.0562)
HS12	0.1951 (0.0008)	1.0000 (0.0000)	-0.0779 (0.1838)	-0.0621 (0.2896)	0.0921 (0.1156)
BH12	0.0800 (0.1723)	-0.0779 (0.1838)	1.0000 (0.0000)	0.0906 (0.1217)	0.0591 (0.3133)
A/IX	-0.0578 (0.3244)	-0.0621 (0.2896)	0.0906 (0.1217)	1.0000 (0.0000)	-0.1383 (0.0178)
A/IN	0.1117 (0.0562)	0.0921 (0.1156)	0.0591 (0.3133)	-0.1383 (0.0178)	1.0000 (0.0000)

TABLE 16: (Continued)

The abbreviations utilized in this matrix are:

BSQ - Buy/Sell Quarterly
HSQ - Hold/Sell Quarterly
BHQ - Buy/Hold Quarterly
BS6 - Buy/Sell 6 Month Test (Etc.)
BS12 - Buy/Sell 12 Month Test (Etc.)

A/IX - Actual vs Index Test
A/IN - Actual vs Initial Holding Test

Tests of Significance by Fund Objective

To test whether a fund's stated objective is related to a manager's ability to select stocks (as highlighted by these tests), the sample was disaggregated according to fund objective. The results of this disaggregation are presented in table (17).

It seems clear that there is no significant difference in prediction skill among the types of funds as "type of fund" was determined here. No Z-statistics are significant and most of the observations are as close to expected as they could possibly be, given the constraint that we must have whole numbers of funds.

Tests of Consistency

The results of the tests of consistency are shown in table (18) below.

There is no evidence of predictability unearthed by this test. None of the variables in the first period have a strong relationship with the same variables in the second period. The R^2 values are very small for all combinations. As was expected, there is a very high correlation among variables within each period.

TABLE 17: TEST OF SIGNIFICANCE BY FUND OBJECTIVE

BUY/SELL (QUARTERLY)

FUND TYPE	# IN SAMPLE	E(+/SIG)	+/SIG	E(-/SIG)	-/SIG	Z
AGG GROWTH	15	1+	2	0	0	0.5
GROWTH	127	12+	13	1+	2	0.33
GRO/INCOME	67	6+	4	1-	0	1.08
INCOME	69	6+	7	1-	1	0.23

HOLD/SELL (QUARTERLY)

FUND TYPE	# IN SAMPLE	E(+/SIG)	+/SIG	E(-/SIG)	-/SIG	Z
AGG GROWTH	15	1+	2	0	0	0.5
GROWTH	127	13-	13	1+	2	0.09
GRO/INCOME	67	7-	6	1-	0	0.28
INCOME	69	7-	7	1-	1	0.08

BUY/HOLD (QUARTERLY)

FUND TYPE	# IN SAMPLE	E(+/SIG)	+/SIG	E(-/SIG)	-/SIG	Z
AGG GROWTH	15	1-	1	0	0	0.53
GROWTH	127	5	5	1+	1	0.22
GRO/INCOME	67	3-	2	1	1	0.52
INCOME	69	3-	3	1+	3	0.15

Key: E(+/SIG) is the number of funds in a given catagory that would be expected by chance to be positive and significant, given the total number of funds that were positive and significant and the percentage of the total funds that is represented by the catagory in question. +/SIG is the actual number of funds in a given catagory that tested positive and significant. The same reasoning applies to the E(-/SIG) and -/SIG headings. Z is the Z statistic calculated as explained in the Statistics section.

TABLE 18: SPEARMAN TEST OF CONSISTENCY

	RANK	RANK1	DIFF	DIFF1	T	T1
RANK	1.00000 (0.0000)	0.05955 (0.3097)	1.00000 (0.0001)	0.05019 (0.3928)	0.95247 (0.0001)	0.06676 (0.2555)
RANK1		1.00000 (0.0000)	0.06168 (0.2951)	1.00000 (0.0001)	0.06419 (0.2785)	0.95704 (0.0001)
DIFF			1.00000 (0.0000)	0.05204 (0.3781)	0.95247 (0.0001)	0.06603 (0.2632)
DIFF1				1.00000 (0.0000	0.06414 (0.2789)	0.95703 (0.0001)
T					1.00000 (0.0000)	0.08699 (0.1415)
T1						1.00000 (0.0000)

The variables "RANK", "DIFF" and "T" represent the first period and the variables "RANK1", "DIFF1" and "T1" represent the second.

TABLE 19: DECILE TEST OF CONSISTENCY

1st HALF RANKING DECILES	2nd HALF RANKING 1	2	3	4	5	6	7	8	9	10	Total
1:	3	3	5	4	2	0	3	1	6	2	29
2:	6	7	2	1	2	1	1	4	1	4	29
3:	4	1	1	1	2	3	3	3	6	5	29
4:	1	3	1	5	2	6	3	3	2	3	29
5:	1	2	8	6	3	3	1	4	1	0	29
6:	2	2	3	2	3	5	3	3	2	4	29
7:	3	2	3	3	4	5	1	1	4	3	29
8:	3	5	1	2	6	1	3	0	2	6	29
9:	4	2	2	4	1	2	2	7	3	2	29
10:	2	2	3	1	4	3	9	3	2	3	32
Tot	29	29	29	29	29	29	29	29	29	32	

Weighted averages (of 2nd half rankings for each 1st half rank):

1	2	3	4	5	6	7	8	9	10
5.27	4.62	6.55	5.83	4.62	5.90	5.48	5.48	5.66	6.00

Sum of 1-5 : 26.89
Sum of 6-10 : 28.52
Sum of 1-3 : 16.44
Sum of 8-10 : 17.14
(Deciles 1-9 contain 29 funds each, decile 10 contains 32)

Table (19) displays little or no clear evidence of predictability. While more of the funds in deciles 1 and 2 (in the first period) do remain in the top half in the second period than migrate to the bottom half (35 to 23), and the funds in deciles 9 and 10 also stay in the bottom half (36 to 25), there is no overwhelming evidence of predictability in this measure. However, there is no evidence here of the negative relationship found in earlier studies.

Summation of Results

The many tests conducted in this study gave new insights into the mutual fund manager stock selection decision process. The study also added to the literature in that it tended to confirm the findings of earlier studies of portfolio managers.

The outcomes of the many tests conducted are as follows:

Buy-Sell-Hold Tests

Returns on stocks bought and on stocks held in general exceeded returns on stocks sold. The results of the overall tests were strong and consistent. The individual fund tests were consistent with the overall tests, but the results were not as strong. These tests support the earlier studies by Mayers and Rice, by Black, by Kon and by Kon and Jen. In addition, there is some evidence that returns on stocks bought exceed those on stocks held.

Actual-Index Test

The results of the overall and the individual fund index tests showed index returns to be greater than actual returns by a wide and significant margin. However, lagged return tests indicated that this result could be due to managers' choosing stocks that had already achieved unusually high returns in the months before their purchase, biasing the index in the direction of higher returns. This test confirmed the often stated but not empirically confirmed assertion that managers do buy stocks after they have achieved unusually high returns.

Actual-Initial Holding Test

The actual-initial test tended to confirm the results of the other tests. Actual returns were significantly higher than the returns that would have been earned had the sample funds simply retained their original holdings.

Other Tests and Correlation

The other tests were not intended to be of major importance in this study, but they did yield some interesting results. As in previous studies, the Sharpe, Treynor and Jensen tests were highly correlated with each other, but they were not, in general, highly correlated with the other tests in this study. Tests to determine whether or not significance was related to fund objective and to examine consistency failed to unearth any significant relationships.

CHAPTER VII

SUMMARY

Final Summation

In final summation, this study tested the buy-sell-hold decisions of the managers of 293 mutual funds. These decisions were tested in a variety of ways with the following results: There is strong evidence that fund managers sell stocks that subsequently earn lower returns than the ones that they either retain or purchase. There is some evidence that there is a difference between the stocks that fund managers buy and the stocks that they retain. Given the biases present, there is no evidence to indicate that fund managers outperform an index composed of the stocks that they have held during the test period. There is evidence that funds outperform an index comprised of each fund's initial holdings. The results of traditional measures in this series of tests appear to mirror the results of previous studies.

Suggestions for Further Research

The opportunities for further research in this area are virtually unlimited. The data set used in this study is new and has not been used by researchers until the last few years. The data is now available at UCLA and the University of Florida and is being utilized by Grinblatt and Titman as well as Radcliffe and Bickum. The methods used in this study may be applied to other data sets, (pension funds, trust funds, insurance companies, managers of other portfolios) and there are many refinements and extentions that can be applied to this methodology. One follow-on study that might be worthwhile would extend the number of observations for individual fund tests (longer time period) to include different economic cycles and increase the probability of detecting significance if indeed it is present. Also, different statistical methods may be applied to the data.

Since the Index test (methodology two) was biased by unusual prior-to-purchase performance, it would be interesting to attempt to remove this bias with a variant of the Copeland and Mayers future benchmark technique. Significant performance relative to an unbiased own portfolio benchmark would strengthen the implications of this study.

While this study has suggested statistically significant benefits in the activities of mutual fund managers, it did not impute or imply economic significance. Further studies might examine whether or not there are gains attributable to the fund managers' activities after transactions costs. The question of whether or not an individual can profit (could have profitted) by following the buy-sell decisions in his/her own portfolio could also be examined.

It is a gross understatement to declare that more definitive information is needed about managed stock portfolios. They are growing at an explosive rate and are the subjects of almost continual ad-hoc and anecdotal news stories. More research is needed in this vital area.

APPENDIX

STATISTICAL METHODOLOGY

STATISTICAL TESTS

1. BUY/SELL/HOLD TESTS

A. Individual fund tests:

(1) Calculated the return for each quarter for stocks bought sold and held by the fund.

(2) Calculated the difference in returns in each quarter (buy-sell, buy-hold, hold-sell).

(3) Tested whether or not difference was statistically different from zero using the Student's T test.

$$T^f = (y^f - \mu)/[\sigma^f/(n^f)]^{1/2}$$

In this series of tests

y^f = the mean difference between (for instance) buys and sells for a given fund,

μ = 0,

σ^f = the standard deviation of the quarterly return differences for each fund,

n^f = the number of quarterly observations (33 for quarterly tests 16 or 15 for six-month tests 8 or 7 for twelve-month tests).

B. Overall tests

(1) Calculated the mean difference for each fund as above.

(2) Tested for significance with the Student's T. (In these tests, n = from 289 to 293 depend--ing on the number of missing observations).

Note: The overall tests were done in this manner for computational tractability [the mainframe (IBM 4341) lacks sufficient memory to compute the results using the complete data sets] and to remove market variation. The samples would have to be assumed to be independent to test all observations at once, and they are matched on time period and manager.

2. ACTUAL/INDEX TESTS

A. Overall test

(1) Calculated the actual and the adjusted index returns for each quarter as detailed in the methodology section.

(2) Multiplied the quarterly returns together to obtain a return relative for the entire period for the actual and the index portfolios.

(3) Subtracted the index return relative from the actual return relative for each fund to obtain the difference.

(4) Tested for significance with the Student's T-test (293 observations).

Note: In this and the overall actual/initial test, the use of the T-test is justified by the central limit theorem. The number of observations needed for this theorem to be effective is, in general, greater than 30. We have 293 observations in this test and 289 in the actual/initial test.

B. Individual fund tests

(1) Calculated the quarterly actual and adjusted index returns as above.

(2) Subtracted the index return from the actual return in each quarter.

(3) Tested whether the differences were significantly different from zero using the Student's T-test (33 observations).

3. ACTUAL/INITIAL TESTS

A. Overall test -- The test was the same as the actual vs index test above, except that no adjustment was needed.

B. Individual fund tests

(1) Calculated the actual and index returns for each quarter.

(2) Multiplied the quarterly returns together to obtain return relatives for the actual and initial returns.

(3) Subtracted the actual from the initial return relative.

(4) Tested for significance using the following T-statistic: $T = (D^f - 0)/Se$

Where D^f is the fund difference found in (3),

Se is the standard error of the overall test.

Note: The individual tests in all cases can be used to confirm the results of the overall tests. With the two sided T-tests done in this study, we would expect to have 2 1/2% of funds outside two standard errors on each side of the distribution. Numbers greater than this on one side or the other would be additional support for the alternative hypotheses. With 293 funds, for instance, we would expect to see about 7 funds, +or-3 on each side. Greater than 13 funds on one side, then would indicate a condition other than chance. (14 +or- 4 on both sides).

4. LAGGED RETURN TESTS

Lagged returns were tested in the same manner as were the overall tests in section 1 above.

5. TEST OF STANDARD DEVIATIONS

A. Calculated the standard deviations of the quarterly returns on the buy, sell and hold portfolios for each fund.

B. Tested whether the difference between the means of the standard deviations were different from zero using a Student's T-test in a paired comparison (buy/sell, buy/hold, hold/sell--293 observations).

6. TESTS OF CORRELATION

Tested for correlation with the Pearson Product Moment correlation.

7. TESTS OF SIGNIFICANCE BY FUND OBJECTIVE

Compared the expected with the actual result by means of the following Z-statistic: $Z = (P_s - P_h)/\sigma P_s$

where $\sigma P_s = (P_s Q_h/n)^{1/2}$,
and P_s is the prior probability,
P_h is the observed outcome,
σP_s is the standard deviation of the prior probability,
Q_h is $1 - P_h$,
n is the number of observed occurances (in this case, the number of significant funds.)

8. TESTS OF CONSISTENCY

A. Rank correlations

(1) The total test period was divided into two periods of 16 quarters (4 years) each. (The 33rd quarter was dropped so that the period lengths would be equal.)

(2) All funds were ranked from 1 to 293 on the difference between the quarterly returns on the fund's purchases and sales.

(3) The funds' rankings in the first period and the second period were compared using Spearman's rank correlation.

B. Test of decile migration

(1) All funds were partitioned into deciles according to the difference between the returns on the funds' purchases and sales.

(2) The funds' decile rankings in the second period were compared to the rankings in the first period.

Note:

Statistical consultants for this dissertation were Dr. Satya N. Mishra, Dr. John J. Lefante and Dr. Robert W. Mee. References were "Mathematical Statistics with Applications" by Mendenhall and Scheaffer, "Statistical Analysis for Business Decisions" by Spurr and Bonini and "Statistics for Modern Business Decisions" by Kyohei Sasaki.

BIBLIOGRAPHY

Arbel, Avner, Steven Carvell and Paul Strebel "Giraffes, Institutions and Neglected Firms" Financial Analysts Journal (May-June 1983), pp.57-63

Arbel, Avner and Bikki Jaggi "Market Information Assimilation Related to Extreme Daily Price Jumps" Financial Analysts Journal (November-December 1982), pp. 60-66

Arbel, Avner and Paul Strebel "Pay Attention to Neglected Firms!" The Journal of Portfolio Management (Winter 1983), pp. 37-42

Arbit, Hal "The Nature of the Game" The Journal of Portfolio Management (Fall 1981), pp. 5-9

Arnott, Robert D. "Cluster Analysis and Stock Price Comovement" Financial Analysts Journal (November-December 1980), pp.56-62

Arnott, Robert D. "What Hath MPT Wrought: Which Risks Reap Rewards?" The Journal of Portfolio Management (Fall 1983), pp.5-11

Ayres, Herbert F. and John Y. Barry "Prologue to a Unified Portfolio Theory" The Journal of Finance (May 1982), pp. 625-635

Babcock, Guilford C. "The Roots of Risk and Return" Financial Analysts Journal (January-February 1980), pp. 56-63

Baesel, Jerome, George Shows and Edward Thorpe "Can Joe Granville Time the Market?" The Journal of Portfolio Management (Spring 1982), pp. 5-9

Banz, Rolf "The Relationship Between Return and Market Value of Common Stocks" Journal of Financial Economics (September 1981), pp. 3-18

Berkman, Neil "Institutional Investors and the Stock Market" New England Economic Review, FRB of Boston (November-December 1977), pp. 188-206

Bernstein, Peter L. "Markowitz Marked to Market" Financial Analysts Journal (January-February 1983), pp. 18-22

Bierman, Harold, Jr. "How Much Diversification is Desirable?" The Journal of Portfolio Management (Fall 1980), pp. 42-44

Bierman, Harold, Jr. "Toward a Constant Price- Earnings Ratio" Financial Analysts Journal (September- October 1982), pp. 62-65

Black, Fischer, "Yes Virginia, There Is Hope: Test of the Value Line Ranking System" Unpublished Manuscript (1971)

Black, Fischer and Moray P. Dewhurst "A New Investment Strategy for Pension Funds" The Journal of Portfolio Management (Summer 1981), pp. 26-34

Black, Fischer, Michael C. Jensen and Myron Scholes "The Capital Asset Pricing Model: Some Empirical Tests, in M. C. Jensen, ed., Studies in the Theory of Capital Markets (1972) Praeger, New York

Bogle, John C. and Jan M. Twardowski "Institutional Investment Performance Compared: Banks, Investment Counselors, Insurance Companies and Mutual Funds" Financial Analysts Journal (January-February 1980), pp. 36-41

Brightman, Jon S. and Barbara L. Haslanger "Past Investment Performance: Seductive but Deceptive" The Journal of Portfolio Management (Winter 1980), pp. 43-45

Brown, Stephen J. and Jerold B. Warner "Measuring Security Price Performance" Journal of Financial Economics (August 1980), pp.205-258

Camp, Robert C. and Arthur A. Eubank, Jr. "The Beta Quotient: A New Measure of Portfolio Risk" The Journal of Portfolio Management (Summer 1981), pp. 53-57

Chance, Don M. "Interest Sensitivity and Dividend Yields" The Journal of Portfolio Management (Winter 1982), pp. 69-75

Chen, Nai-Fu, Thomas E. Copeland and David Mayers "A Comparison of APM, CAPM and Market-Model Portfolio Performance Methodologies: The Value Line Case (1965-1978)" Unpublished Manuscript, University of Chicago, University of California, Los Angeles (1983)

Coggin Daniel and John E. Hunter "Problems in Measuring the Quality of Investment Information: The Perils of the Information Coefficient" Financial Analysts Journal (May-June 1983), pp. 25-33

Condon, Kathleen A. "Measuring Equity Transactions Costs" Financial Analysts Journal (September-October 1981), pp.57-60

Copeland, Thomas E. and David Mayers "The Value Line Enigma (1965-1978)" Journal of Financial Economics (October 1982), pp. 289-321

Cornell, Bradford "Asymmetric Information and Portfolio Performance Measurement" Journal of Financial Economics (July 1979), pp. 381-399

Dunn, Patricia C. and Rolf D Theisen "How Consistently Do Active Managers Win?" The Journal of Portfolio Management (Summer 1983), pp. 47-50

Dybvig, Philip H. "Some New Tools for Testing Market Efficiency" Unpublished Manuscript, Yale School of Organization and Management (1983)

Dybvig, Philip H. and Stephen A. Ross "The Simple Analytics of Performance Measurement" Unpublished Manuscript (October 1980)

Edmister, Robert O. and James B. Greene "Performance of Super-Low-Price Stocks" The Journal of Portfolio Management (Fall 1980), pp. 36-41

Fama, Eugene F. "Efficient Capital Markets: A Review of Theory and Empirical Work" The Journal of Finance (May 1970), pp. 383-417

Fama, Eugene F. "Components of Investment Performance" The Journal of Finance (June 1972), pp. 551-567

Ferguson, Robert "Performance Measurement Doesn't Make Sense" Financial Analysts Journal (May-June 1980), pp. 59-6269

Fielitz, Bruce D. and Myron T. Greene "Shortcomings in Portfolio Evaluation via MPT" The Journal of Portfolio Management (Summer 1980), pp. 13-19

Firth, Michael A. "The Investment Performance of Unit Trusts in the Period 1965-75" The Journal of Money Credit and Banking (November 1977), pp. 597-604

Friend, Irwin and Marshall Blume "Measurement of Portfolio Performance Under Uncertainty" The American Economic Review (1970), pp. 561-575

Gaumnitz, Jack E. "Appraising Performance of Investment Portfolios" The Journal of Finance (1970), pp. 555-560

Goodman, David A. and John W. Peavy, III "Industry Relative Price-Earnings Ratios as Indicators of Investment Returns" Financial Analysts Journal (July-August 1983), pp. 60-66

Graver, Robert R. and Nils H. Hakanssen "Higher Return, Lower Risk: Historical Returns on Long-Run Actively Managed Portfolios of Stocks, Bonds and Bills, 1936-1978" Financial Analysts Journal (March-April 1982), pp. 39-53

Grinblatt, Mark "How To Evaluate A Portfolio Manager" Unpublished Manuscript, University of California at Los Angeles (1986)

Grinblatt, Mark and Sheridan Titman "A Comparison of Measures of Abnormal Performance on a Sample of Monthly Mutual Fund Returns" Unpublished Manuscript, University of California at Los Angeles (1986)

Groth, John C., Wilber G. Lewellen, Gary G. Schlarbaum and Ronald C. Lease "Security Analysts: Some Are More Equal" The Journal of Portfolio Management (Spring 1978), pp. 43-48

Harrington, Diana R. "Whose Beta Is Best?" Financial Analysts Journal (July-August 1983), pp. 67-73

Hawawini, Gabriel "Why Beta Shifts as the Return Interval Changes" Financial Analysts Journal (May-June 1983), pp. 73-77

Hawawini, Gabriel A. and Ashok Vora "The CAPM and the Investment Horizon: Comment" The Journal of Portfolio Management (Fall 1982), pp. 66-70

Hayek, F. A. "The Use of Knowledge in Society" The American Economic Review (September 1945), pp. 519-530

Hill, Joanne M. "Is Optimal Portfolio Management Worth the Candle?" The Journal of Portfolio Management (Summer 1981), pp. 59-69

Hitschler, Anthony W. "To Know What We Don't Know or The Caribou Weren't in the Estimates" Financial Analysts Journal (January-February 1980), pp. 28-32

Ibbotson, Roger G., Richard C. Carr and Anthony W. Robinson "International Equity and Bond Returns" Financial Analysts Journal (July-August 1982), pp. 34-56

Ibbotson, Roger G., Jeffrey J. Diermeire and Laurence B. Siegel "The Demand for Capital Market Returns: A New Equilibrium Theory" Financial Analysts Journal (January-February 1984), pp.22-33

Ibbotson, Roger G. and Laurence B. Siegel "The World Market Wealth Portfolio" The Journal of Portfolio Management (Winter 1983), pp. 5-17

Jaffe, Jeffrey F. "Special Information and Insider Trading" The Journal of Business (1973), pp. 410-428

Jensen, Michael C. "The Performance of Mutual Funds in the Period 1945-1964" Journal of Finance (May 1968), pp. 389-416

Jensen, Michael C. "Risk, the Pricing of Capital Assets, and the Evaluation of Investment Portfolios" Journal of Business (April 1969), pp. 167-247

Jensen, Michael C. "Capital Markets: Theory and Evidence" Bell Journal of Economics and Management Science (1972), pp. 357-398

Jobson, J. D. and Bob Korkie "Putting Markowitz Theory to Work" The Journal of Portfolio Management (Summer 1981), pp.70-74

Jobson, J. D. and Bob Korkie "Performance Hypothesis Testing With the Sharpe and Treynor Measures" The Journal of Finance (September 1981), pp. 889-908

Jobson, J. D. and Bob Korkie "Potential Performance and Tests of Portfolio Efficiency" Journal of Financial Economics (October 1982), pp. 433-466

Jobson, J. D. and Bob Korkie "On the Jensen Measure and Marginal Improvements in Portfolio Performance: A Note" The Journal of Finance (March 1984), pp. 245-251

Johnson, James M. and Howard P Lanser "Dividend Risk Measurement and Test of the CAPM" The Journal of Portfolio Management (Winter 1981), pp. 50-54

Kon, Stanley J. "The Market Timing Performance of Mutual Fund Managers " The Journal of Business (1983), pp.323-347

Kon, Stanley J. and Frank C. Jen "The Investment Performance of Mutual Funds: An Empirical Investigation of Timing, Selectivity and Market Efficiency" Unpublished Manuscript (1977)

Korkie, Bob "External vs. Internal Performance Evaluation" The Journal of Portfolio Management (Spring 1983), pp. 36-41

Krasner, William S. "New Approaches to Portfolio Theory and Stock Market Equilibrium--Minimax Behavior in Portfolio Selection" The Journal of Finance (May 1982), pp. 609-614

Kritzman, Mark "Can Bond Managers Perform Consistently?" The Journal of Portfolio Management (Summer 1983), pp. 54-56

LeBaron, Dean "Reflections on Market Inefficiency" Financial Analysts Journal (May-June 1983), pp. 16-23

Lerner, Eugene M. and Pochara Theerathorn "The Returns of Different Investment Strategies" The Journal of Portfolio Management (Summer 1983), pp. 26-28

Levy, Haim "Portfolio Performance and the Investment Horizon" Management Science (August 1972), pp. B645-B653

Levy, Haim "The CAPM and the Investment Horizon" The Journal of Portfolio Management (Winter 1981), pp.32-40

Levy, Haim "A Test of the CAPM via a Confidence Level Approach" The Journal of Portfolio Management (Fall 1982), pp. 56-61

Litzenberger, Robert H. and Krishna Ramaswamy "The Effect of Personal Taxes and Dividends on Capital Asset Prices" Journal of Financial Economics (March 1979), pp. 163-195

Loeb, Thomas F. "Trading Cost: The Critical Link Between Investment Information and Results" Financial Analysts Journal (May-June 1983), pp. 39-44

Lustig, Ivan L. and Philip A. Leinbach "The Small Firm Effect" Financial Analysts Journal (May-June 1983), pp. 46-49

Markowitz, Harry "Portfolio Selection" The Journal of Finance (March 1952), pp. 77-91

Mayers, David and Edward M. Rice "Measuring Portfolio Performance and the Empirical Content of Asset Pricing Models" Journal of Financial Economics (March 1979), pp. 3-28

McDonald, John G. "Objectives and Performance of Mutual Funds, 1960-1969" Journal of Financial and Quantitative Analysis (June 1974), pp. 311-333

McElreath, Robert B. Jr. and C. Donald Wiggins "Using the COMPUSTAT Tapes in Financial Research: Problems and Solutions" Financial Analysts Journal (January-February 1984), pp. 71-75

Mennis, Edmund A. "A New Method for Evaluating Pension Fund Portfolios" The Journal of Portfolio Management (Winter 1979), pp. 34-40

Merrett, A. J. and Gerald D. Newbould "Integrating Financial Performance and Stock Valuation" The Journal of Portfolio Management (Fall 1983), pp. 27-32

Merton, Robert C. "On Estimating the Expected Return on the Market" Journal of Financial Economics (August 1980), pp. 323-361

Miller, Edward M. Jr. "How to Win at a Losers Game" The Journal of Portfolio Management (Fall 1978), pp.17-24

Miller, Merton H. and Myron Scholes "Rates of Return in Relation to Risk: A Reexamination of Some Recent Findings" Studies in the Theory of Capital Markets (New York: Praeger Publishers 1972), pp. 47-78

Miller, Paul F., Jr. "Managing Investment Portfolios: A Revision Article" The Journal of Portfolio Management (Winter 1983), p. 57

Milne, Frank and Ross L. Watts " Corporate Information: A Private or Public Good?" Unpublished Manuscript, University of Rochester (August 1977)

Murphy, J. Michael "Why No One Can Tell Who's Winning" Financial Analysts Journal (May - June 1980), pp. 49-57

Nagorniak, John "Risk Adjusted Equity Performance Measurement" The Journal of Finance (September 1980), pp. 555-561

Olsen, Robert A. "Sample Size and Markowitz Diversification" The Journal of Portfolio Management (Fall 1983), pp. 18-21

Prakash, Prem and Alfred Rappaport "Information Inductance and its Significance for Accounting" Accounting Organizations and Society (1977), pp. 29-38

Pulliam, Kenneth P. "How to Produce Value from Portfolio Measurement" The Journal of Portfolio Management (Fall 1981), pp. 13-16

Reinganum, Marc R. "Abnormal Returns in Small Firm Portfolios" Financial Analysts Journal (March-April 1981), pp. 52-56

Reinganum, Marc R. "Misspecification of Capital Asset Pricing" Journal of Financial Economics (September 1981), pp. 19-46

Roll, Richard "A Critique of the Asset Pricing Theory Tests" Journal of Financial Economics (April 1977), pp. 129-175

Roll, Richard "Ambiguity When Performance is Measured by the Securities Market Line" The Journal of Finance (September 1978), pp.1051-1069

Roll, Richard "The Expected Return Benchmark and Portfolio Performance Evaluation" Unpublished Manuscript, University of California Los Angeles (April 1980)

Roll, Richard "Performance Evaluation and Benchmark Errors" Unpublished Manuscript, University of California Los Angeles (November 1980)

Roll, Richard "A Possible Explanation of the Small Firm Effect" The Journal of Finance (September 1981), pp. 879-888

Roll, Richard "On Computing Mean Returns and the Small Firm Premium" Unpublished Manuscript, University of California Los Angeles (May 1983)

Roll, Richard "Vas Ist Das?" The Journal of Portfolio Management (Winter 1983), pp. 18-28

Rosenberg, Barr "The Capital Asset Pricing Model and the Market Model" The Journal of Portfolio Management (Winter 1981), pp.5-16

Rosenberg, Barr "The Current State and Future of Investment Research" Financial Analysts Journal (January-February 1982), pp.43-50

Rosenberg, Barr and Andrew Rudd "Return Expectations in Active Investment Management" The Journal of Finance (May 1982), pp.543-554

Sharpe, William F. "Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk" The Journal of Finance (September 1964), pp. 425-442

Sharpe, William F. "Mutual Fund Performance" Journal of Business (January 1966), pp.119-138

Shawky, Harry A. "An Update on Mutual Funds: Better Grades" The Journal of Portfolio Management (Winter 1982), pp.29-34

Sherman, Eugene J. "A Gold Pricing Model" The Journal of Portfolio Management (Spring 1983), pp.68-70

Smidt, Seymour "Investment Horizons and Performance Measurement" The Journal of Portfolio Management (Winter 1978), pp.18-22

Smith, Gary "A Simple Method for Estimating Intrinsic Value" The Journal of Portfolio Management (Summer 1982), pp. 46-49

Stock, Duane "Does Active Management of Municipal Bond Portfolios Pay?" The Journal of Portfolio Management (Winter 1982), pp. 51-55

Theobald, Michael "On Estimating Betas that change" The Journal of Portfolio Management (Fall 1982), pp. 62-65

Treynor, Jack L. "How to Rate Management of Investment Funds" Harvard Business Review (January-February 1965), pp.63-75

Unger, Raymond F. "Using Outside Research" The Journal of Portfolio Management (Spring 1983), pp. 43-45

Vandell, Robert F. "Is Beta a Useful Measure of Security Risk?" The Journal of Portfolio Management (Winter 1981), pp. 23-31

Vandell, Robert F. and Mark T. Finn "Portfolio Objective: Win Big, Lose Little!" The Journal of Portfolio Management (Summer 1982), pp. 37-45

Veit, E. Thoedore and John M. Cheney "Are Mutual Funds Market Timers?" The Journal of Portfolio Management (Winter 1982), pp. 35-42

Walking, Ralph A. and Robert O. Edmister "Are There Commission Cost Side-Effects from Portfolio Management Decisions?" Financial Analysts Journal (July-August 1983), pp.52-59

Wallace, Anise "Is Beta Dead?" Institutional Investor (July 1980), pp. 23-30

Weston, J. Fred "Developments in Financial Theory" Financial Management Tenth Anniversary Issue (1981), pp. 5-22

Wilcox, Jarrod W. "The P/B--ROE Valuation Model" Financial Analysts Journal (January-February 1984), pp. 58-66

Zeckhauser, Richard and Victor Niederhoffer "The Performance of Market Index Futures Contracts" Financial Analysts Journal (January-February 1983), pp. 59-65

BIOGRAPHICAL SKETCH

Gilbert W. Bickum grew up in Newton, Mass., and attended Newton Public Schools. His undergraduate degree is in mechanical engineering from Tufts University in Medford, Mass. While Serving in the U.S. Navy, he completed an MBA at the University of North Florida in Jacksonville. After retiring from the Navy, he completed his Ph.D. at the University of Florida, concentrating in finance. He taught finance and investment courses at the University of South Alabama from 1984 through 1987, and is currently an Assistant Professor of Finance at the University of Eastern Kentucky.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adaquate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Robert C Radcliffe

Robert C. Radcliffe, Chairman
Associate Professor of Finance,
Insurance and Real Estate

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adaquate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

A. R. Abdel-Khalik

Rashad A. Abdel-khalik
Graduate Research Professor of
Accounting

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adaquate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Roger Huang

Roger Huang
Associate Professor of Finance,
Insurance and Real Estate

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adaquate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

T. Craig Tapley

T. Craig Tapley
Assistant Professor of Finance,
Insurance and Real Estate

This dissertation was submitted to the Graduate Faculty of the Department of Finance, Insurance, and Real Estate in the College of Business Administration and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

December 1987

Dean, Graduate School

UNIVERSITY OF FLORIDA
3 1262 08285 309 3